Instituto Tecnológico y de Estudios Superiores de Monterrey

EGADE Business School

Ciudad de México

Markets never give in: an asset price bubble analysis

A dissertation presented by

Carlos Armando Franco Ruiz

Submitted to the
EGADE Business School
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Financial Sciences

Mexico City, Mexico, November 6, 2020

Dedication

For my parents and brother,

with all my love.

Acknowledgements

After a road with ups and downs that test our knowledge, perseverance, and goals, it finally concludes. Still, in reality, it is just the beginning of the path of a researcher.

The Ph.D. has represented one of the most significant challenges, and without a doubt, one of the best experiences that life has allowed us to face. The search for academic contributions, such as a Ph.D., is not a hundred-meter race but a marathon in all senses. As with any career, none runs alone, and therefore I should thank all the people who made the completion of this Ph.D. possible. I hope to be able to mention each of them, and in the case of having omitted someone, I gratitude all in advance. I will remember you at all times in my mind and my heart.

First of all, I want to acknowledge Dr. Daniel Cerecedo Hernández for giving me the benefit of the doubt, agreeing to be my tutor, and giving us the freedom to focus our research on the area of our interest, which in my case, were the financial bubbles. Similarly, I want to thank Dr. José Antonio Núñez Mora for being principally our academic guide, somewhat similar to our "father of research." A part of this work is due to their support, both academically, personally, and professionally. Thanks to these two people again.

Likewise, I want to thank Dr. Guillermo Benavides Perales for having accepted to be one of my committees, offering us his knowledge in each of his lessons on time series and showing us that with effort and dedication, the combination of the academy with the industry is possible. I want to thank Dr. José Juan Chávez Gudiño for all the Saturday lectures that represented a challenge within risk management and instilled that we impose ourselves all the limits. I want to thank Dr. Montserrat Reyna, Dr. Carlos Viguri, and Dr. Alberto Ortiz Bolaños for their constant intellectual challenges.

On the other hand, I want to thank my friends and colleagues Mario Iván Contreras, Alfredo Ramírez, Eduardo Sánchez, and Breno Madero, to contribute to building a virtuous circle throughout the Ph.D.

Secondly, I want to thank the Celorio family (Jesús Agustín, Dulce María, José Agustín, and Alexandro) for giving me their professional support and friendship. In the same way, I want to thank my parents and brother for all the effort made. This accomplishment is for them. I want to offer special recognition to Pamela Moncayo for having been a meaningful part of this thesis's development in its proof-reading and the contributions made for the completion of this research. Again, thanks for everything.

Finally, I want to thank Dr. José Ernesto Amorós and his team for their commitment to improving this Ph.D. I want to offer infinite gratitude to the Tecnológico de Monterrey for supporting me with an academic scholarship throughout my Ph. D. and for having been my alma mater for more than a decade.

"Success is not final, failure is not fatal; it is the courage to continue that counts."

- Winston Churchill

Markets never give in: an asset price bubble analysis

By

Carlos Armando Franco Ruiz

Abstract

This thesis aims to analyze asset price bubbles, where we developed, in Chapter I, a brief historical crashes description and a bibliometric analysis of 2,494 articles. In Chapter II, we studied the presence of financial bubbles in fifty stocks that constitute the S&P 500 index, using the generalized augmented Dickey-Fuller (GSADF) test proposed by Phillips et al. (2011, 2015). We found one hundred six bubbles in fifteen assets and detected that in the last decade (2010-2020), there is an increasing pace of this phenomenon.

In Chapter III, we developed the ability of the Normal Inverse Gaussian distribution (NIG) to fit the returns of eight stocks where we found in the previous chapter at least one bubble-type behavior in the period from January 3, 2000, to December 31, 2009 (1P), and from January 4, 2010, to April 29, 2020 (2P). For the first period, the NIG could fit the mentioned segment; therefore, we estimate at different levels of confidence the VaR and CVaR for the in-sample-data (1P). We took the maximum expected loss and shortfall values and applied them to the out-of-the-sample (2P). In conclusion, we obtained a good adjustment to the second period (2P) and found the NIG differences compared to the Generalized Hyperbolic (GH) are just marginal. At the same time, we benefit the NIG is close under convolution and minor computational effort evaluation.

In Chapter IV, we implemented a model-based clustering method of the Gaussian mixture model to categorize previously identified asset price bubbles and three dropdown scenarios of the S&P 500 index for 2020. We took an approach based on the price-driven identification: bubble size and crash size. We obtained different Gaussian cluster models and concluded that the Gaussian mixture model is a gold standard for further investigations. Finally, in Chapter V, we developed the previous chapters' final remarks that include all supervisors' valuable feedback.

CONTENTS

List of tables

List of figures

1. INTRODUCTION

"I can calculate the movement of the stars, but not the madness of men".[1] According to the definition of the *Real Academia Española*, greed is understood as the excessive desire for wealth or vehement desire for some good things. On the other hand, fear is anguish over a real or imaginary risk, damage, or suspicion or apprehension that someone has that something opposite happens to what they want. This conceptualization is the economic agents' behavior when they buy high and sell low, once euphoria is exhausted and madness vanishes.

Financial bubbles embody and highlight the holy grail or the long-awaited opportunity of a free lunch in the financial market's discipline. Non a simple task, "no warning can save a people determined to grow suddenly rich"[2]. Humans have not learned from their mistakes throughout history, and it is a lesson that they should understand. In the 21st century, in recent economic history, this lecture has been repeated more frequently than ever, the Dotcom crisis, the 2008 crisis, the explosion in the Bitcoin's price, and for more than a decade a steady increase in general in major stock indexes since 2009.

However, we must not forget the systemic study of the boom and crash of the so-called Tulip Mania, South Sea Bubble, the Great Depression, among others, that could function as fingerprints or keys for future events around the corner.

1.1 MOTIVATION

Asset price bubbles have been a rude awakening for investors, central bankers, and policymakers due to their unpredictability and fascination, yet these great shocks are often misunderstood. There is a misconception that these events are unpredicted; nevertheless, we can keep providing information on the existing complex system that

[1] Quote attributed to Sir Isaac Newton in his participation in the financial bubble of the South Sea Company of 1720.
[2] Quote attributed to Samuel Jones Loyd, known as Lord Overstone.

financial time series and, in particular, exuberant episodes represent, changing in some sense the paradigm between the academic and the practitioner perspectives.

1.2 PROBLEM STATEMENT AND CONTEXT

Asset price bubbles are as old as human greed; nevertheless, this concept has gained popularity, posing the need to study these exuberant periods. They can affect the entire financial markets systematically, causing losses of trillions of dollars worldwide of the stock market capitalization.

One major issue regarding the stylized facts of financial assets series is the continuous adoption of the normal distribution to fit these times series. The problem lies within the nature of financial assets, and it increases when we are in the presence of a bubble-type behavior, where fat-tails are not considered in risk management.

Several academics have made a wide range of attempts to capture the exuberant nature of financial series and identify bubbles to neutralize its harmful effects. Nonetheless, there is not a standard approach to foresee the start and end of such phenomena. Hence, we established the main objectives of this investigation:

1. Encourage the detection of financial bubbles.
2. Consider the ability of non-normal distributions to fit time-series returns.
3. Categorize financial bubbles according to certain factors.

1.3 RESEARCH QUESTION

How would financial practitioners identify, fit, and categorize asset price bubbles in complex past, present, and future financial markets?

1.4 SOLUTION OVERVIEW

We must comprehend the essential characteristics of financial assets such as they possess heavy-tails; if not, we will not recognize asset price bubbles' features. Consequently, in this thesis, we proposed in first place for market excesses the application of a well-known identification test developed by Phillips et al. (2011, 2015). Secondly, we suggest the implementation of the NIG distribution to fit bubble-type behavior periods. Thirdly, we proposed a categorization of financial bubbles according to the bubble and crash size.

1.5 DISSERTATION STRUCTURE

Chapter I presents a description of the concept of a financial bubble and a brief historical crashes explanation from Tulip Mania, passing through South Sea Company, the Great Crash of 1929, the Dotcom bubble, the Great Financial crisis, and the Bitcoin crash, to name a few of the remarkable dropdowns humankind has perceived. Finally, we developed a state-of-the-art approach by using a bibliometric analysis of 2,494 articles.

In Chapter II, we studied fifty shares that constitute the S&P 500 index, where we applied the GSADF test to identify multiple bubbles in these assets. From this test, we detected 106 bubbles in 15 assets.

On the other hand, in Chapter III, we took the 15 assets from the previous chapter, made a smaller sample according to specific criteria, and proved that the NIG distribution could fit episodes that contain bubble-type behaviors. Therefore, we estimated with the NIG and GH distribution different VaR and CVaR levels of the in-sample data and applied to the out-of-the-sample period to observe how they adjust.

In Chapter IV, we create an asset price bubble categorization through two concepts, bubble size and crash size, by applying a Gaussian finite mixture model to 17 explosive behaviors previously identified. Furthermore, we set three possible fall scenarios in 2020 of the S&P 500 index, and show that the analysis of financial bubbles is a complex dynamic movement.

Finally, in Chapter V, we mentioned the final remarks that, from our point of view, could help the lector to go deep in the information about financial bubbles in order to fulfill the suggestion of all supervisors.

2. CHAPTER I: UPS AND DOWNS OF THE FINANCIAL MARKETS

2.1 WHAT IS A BUBBLE?

The initial question would be what is the definition of a “bubble” or what we should understand by this concept. Bubbles have existed throughout history with abnormal moments of euphoria and panic. Aliber and Kindleberger (2017) defined it as “a generic term for the increases in the prices of securities or currencies in the mania phase of the cycle that cannot be explained by the changes in the economic fundamentals” (p. 21). Durlauf and Blume (2008) described bubbles as a dramatic price increase followed by collapse or when the price exceeds the asset's fundamental value. This exuberance in prices usually is the precedent of a crisis.

Tirole (1982) stated that an asset's price is the sum of the fundamental and speculative value. The uncertain value is what he called the “price bubble” (p. 1,172). Likewise, Blanchard and Watson (1982) indicated that the price is the sum of two components, the fundamental value, and the bubble, but as “bubbles can take many forms,” a general classification for its discovery is complicated (p. 13). Whereas Flood and Garber (1980) specified that “a price bubble exists when the expected rate of market price change is an important factor determining current market price,” and they pointed out that given their presence, the explanation of their path and termination is of utmost importance (p. 745 - 760).

Evans (1991) described a bubble as an eruption that grows in a specific phase and then collapses. West (1987) refers to speculative bubbles as self-fulfilling rumors of potential stock price fluctuations that later drop (“popping”), reflecting a market overreaction (p. 553-559). Allen and Gale (2000) typically punctuate asset price bubbles in three phases in a more recent approach. The first is financial liberalization, where central banks increase loans or carry out other similar activities accompanied by a continuous increase in assets that can last for several years. The second phase is when the bubble bursts where the fall periods go from a few days to more extended periods. Finally, the last contemplates the default of different economic agents who borrowed to buy assets with inflated prices.

Consequently, we note that convergence to a single definition is not yet stipulated; however, we could generally define it as an ascending acceleration above an asset's fundamental value. To provide insight or overview applied to the definitions provided above, we will show some examples of historical crashes in the next section.

2.2 BRIEF HISTORICAL CRASHES

After the second collapse during the 21st century in the years to come, the prosperity of continued growth in financial markets reborn. Since the S&P 500 index's highest price during the 2008 crisis, around 1,565.15 on October 9, 2007, it has increased 116.35 percent to 3,386.15 on February 19, 2020. Nevertheless, the fear of a market collapse awoke after a pandemic broke out, where the S&P 500 index fell almost 32 percent in a month. In Fig. 1, we exposed a calendar heat map of the previous stock index's daily returns from 2015 to April 17, 2020, where we noticed the collapse in this period.

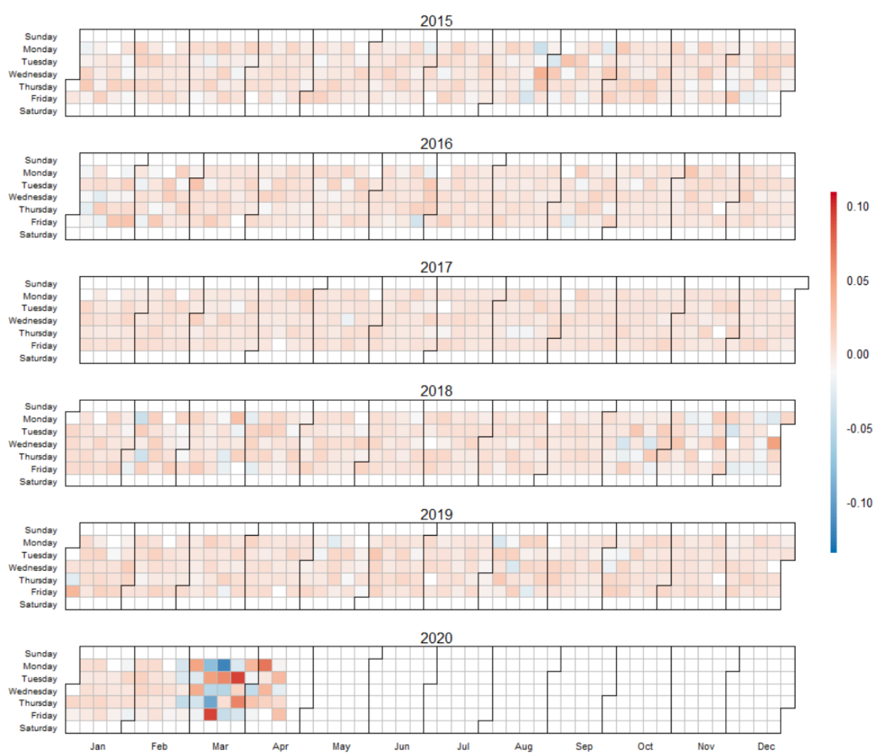

Fig. 1 Calendar heat map of the daily returns of the S&P500 index from 2015 to April, 17, 2020. The author created the image.

So, once again, investors notice that securities prices could not rise forever. The speculative frenzy of a bubble comes from the birth of financial markets. People see in a “boom” unlimited potential returns, but when the “bust” appears, people panic, and their gains dissipate. The previous scenario is not new; it has repeated in many historical financial crashes. As Horace (Quintus Horatius Flaccus) said: "A heart well prepared for adversity in bad times hopes, and in good times fears for a change in fortune."

2.2.1 THE TULIP MANIA

The Dutch created a thriving economy at the beginning of the sixteenth century, where the Netherlands was in a period of great prosperity and a driving force for a newly-created global economy (Day, 2004). Amsterdam became the top commercial emporium and represented the center of the trade of the northwestern part of Europe (Sornette, 2017). In this framework, aimed at an environment where economic agents sought the accumulation of wealth and cultivated expensive tastes, caused the emergence of the tulip futures market.

Consequently, tulips represented a luxury asset and, therefore, the participants' wealth in this market. Speculation on the tulip bulbs represented at that historical moment a safe investment. By the year 1634, buying tulip futures contracts in fall and winter for future delivery was ongoing practice. During most of the tulip speculation, high prices centered on bulbs with a mutation called "breaks," caused by a virus. Nevertheless, the premium for the ultra-rare breaks did not last long as new varieties depressed the prices for older ones (Day, 2004, p. 151-163).

Speculation intensified due to the frenzy of what the Dutch called “windhandel,” which is the buying and selling of futures without the goods’ current possession (Sarna, 2010). These relationships were prohibited, but a continuous informal market was on par (Day, 2004). Finally, tulip prices collapsed in February 1637 when “the contract price of tulips … reached a level that was about 20 times higher than in both early November

1636 and early May 1637" (Thompson, 2007, p. 99). It is worth mentioning that Thompson establishes a different point of view on tulipmania (2007).

2.2.2 SOUTH SEA COMPANY

The South Sea Company is the first stock market crash in England in 1720. Schachter, Gerin, Hood, and Anderassen (1985) argued that it is "the villain of the colossal scam," where the creation of the Company emerged in May 1711, by an Act of Parliament which offered a promising solution for the National Debt. This scheme contemplated that short-term public debt holders could exchange debt instruments for shares of the South Sea Company. The government-guaranteed an annuity to this Company, which permitted a 5% interest on the stock. As an incentive for investors, England granted the monopoly of trade to South America (Spanish ports). However, during 1711 and 1720, Spain and England were in a period of war or continuous hostility, so the "monopoly was at best a chimera." Nevertheless, these measures produced interest for speculators and investors of those times (p. 324).

This bubble culminated in a story of mass frenzy, political corruption, and public upheaval (Sornette, 2017). To understand the political and economic context, it is necessary to know that England wanted to vent its national debt in those moments of high tension. "… for its time, [it was] an intricate piece of financial engineering. The government's benefits were plain: they could dispense with the irredeemables, reduce interest rates on the National Debt, and earn £7.5 million. The benefits to the Company were more opaque [sic]" (Deringer, 2015, p. 652).

The South Sea Company scheme came to an end when the Sword Blade Company failed under the Bank of England's demanding redemption of its notes or their payment in specie on September 24. There is uncertainty about whether the Bank of England terminated the scheme; however, it is clear that it denoted that credit was restricted (Neal, 1993). Notwithstanding, the chaos began since the Parliament awarded the contract of assuming most of the National Debt at the beginning of February 1720 (Schachter et al., 1985).

2.2.3 THE GREAT CRASH OF 1929

After almost a century, we continue to gather information from the fall of October 29, 1929, where investors (as in any other time) had a belief that assumed unlimited horizons. For the general public, the Great Crash and the Great Depression are synonymous; however, they do not make the same cataclysm for economists. The first refers to the decline in stock prices in October 1929, while the second involves a tremendous drop in real output between 1929 and 1933 (Romer, 1990). There are different positions regarding this dichotomy, and in particular, on what variables could have caused the Great Recession, as Friedman and Schwartz pointed out in “A monetary history of the United States, 1867-1960,” or Bernanke in “Irreversibility, uncertainty, and cyclical investment,” to name a few. In this order of ideas, Romer raises an “uncertainty hypothesis,” where she investigated that the extreme variability of stock prices caused uncertainty about people’s future income levels and, therefore, postponed the purchase of the irreversible durable goods. This view is relevant because it exemplifies the difference between the collapses of 1929 and 1987, where consumers perceived the latter as a one-time aberration (1990, p. 599-602).

The boom of the 1920s headed to a bubble (Ribstein, 2003). After World War I and a post-war recession, from 1922 to 1929, the GNP grew at an annual average of 4.7 percent, while unemployment remained at an average of 3.7 percent (White, 1990). Heilbroner and Milberg (2008) comment that few Americans were approaching an economic calamity. The rise of the stock market had attracted (according to these authors), perhaps 10 million investors who perceived an increase in their capital without any effort. Speculation was in the environment; however, the risk was justifiable. So that a person who invested a thousand dollars a year since 1921 in some representative stock, by 1925, would have had six thousand dollars and by 1928 would have reached twenty thousand dollars. It is worth mentioning that this growth trend was only the beginning since by mid-1929, industrial shares' average increased almost the same as the previous year. For the eighth month of 1929, the profits obtained during the summer period exceeded the returns of the entire year of 1928. On Tuesday, October 29, 1929, everything fell; for example, at the end of the session on that day, investors sold 16,410,000 shares while thirty billion dollars vanished in just a few weeks. By 1933, one out of every two dollars disappeared from final production.

2.2.4 DOTCOM BUBBLE

The wealth potential in a new market that causes a bubble in some specific sector is not new. Occasionally, the appearance of bubbles in the market represents a bleak outlook for investors and workers. However, opportunities arise that can make companies that survive after the bubble collapse become profitable economic agents. In this order of ideas, before the beginning of the dotcom bubble, the internet's rapid growth attracted interest from investors and entrepreneurs (Panko, 2008).

Before the new millennium, the previous decade consisted of seeking market leadership that used the internet as a profit generator, and not as a complementary tool to its core business. Within this bubble, Razi, Tarn, and Siddiqui (2004) divided them into two categories, controllable and uncontrollable causes. The first one refers to strategic, operational, and technical reasons, while the latter likewise has technical and behavioral causes. These authors contemplated that the bubbles' causes are not attributable to a single factor but rather to a combination of variables, such as the behavior of a complex system.

The "prediction of the financial pundits" of the dotcom bubble became true wherein a short time, the surprising growth of several stock prices during the 1990s finished and collapsed (Bose and Pal, 2006, p. 960). In a broader explanation Kaizoji, Leiss, Saichev, and Sornette (2015) mentioned that "the valuation of the Internet stock index went from a reference value 100 in January 1998 to a peak of 1,400.06 in March 9, 2000, corresponding to an annualized return of more than 350%! A year and a half later, the

Internet stock valuation was back at its pre-1998 level" (p. 290). In addition to the above, by February 2000, "the Internet sector equaled 6 percent of the market capitalization of all US public companies and 20 percent of all publicly traded equity volume" (Ofek and Richardson, 2003, p. 1,113).

2.2.5 THE GREAT FINANCIAL CRISIS

The 2008 crisis turned out to be a more complex credit crunch than the previous problems that the markets have had, because this time, financial innovation allowed an alternative to packaging and reselling assets. This situation paired with the offer of mortgages to individuals lacking the adequate economic and credit profile; however, the 2008 crisis originated due to the risk mispricing of these products. These new financial products consisted of subprime and other types of mortgages, later sold as sophisticated instruments. Despite the composition of these assets, the rating agencies gave high ratings, which would mean that they were considered safe. However, the framework that caused assets to be related based on the movement of house prices caused foreclosures on mortgages to increase due to the fall in these prices (Mizen, 2008).

Yeo (2010) mentioned that the 2008 global financial crisis is often referred to as the largest socio-economic-political event since the 1950s. He pointed out that the evolutionary process of the collapse of financial markets developed in four stages, possibly with an additional fifth stage. The first stage began in the first month of 2007 when financial institutions reported losses caused to mortgage defaults. The second stage joined the problem that was already happening to the United States. The rapid fall in houses' value within a slow economic scenario caused catastrophic damage to the property markets in the United Kingdom and other parts of the European Union. Phase three started in January 2008, where the global banking credit spreads for AA-rated companies widened over 175 points, causing a shutdown of the asset securitization markets. After the rescue of Bear Sterns and its absorption by JP Morgan-Chase, the crisis' effects had already spread, opening the fourth stage. Since September 2008, several institutions faced a bailed out; however, when US regulators failed to rescue

Lehman, the problem took over the world due to the importance and relationship that this institution represented in the mortgage market and insurance. In an equally discouraging context, the United States and other leading economies went into a technical recession, resulting in a global recession. Finally, for the last phase, around the third quarter of 2009, this author states that the United States and, to a certain extent, the United Kingdom, are in the process of stabilizing their economies.

2.2.6 BITCOIN CRASH

In a scenario where financial transactions were only carried out almost exclusively through trust third parties, where there were certain deficiencies such as 1) weaknesses of the trust-based model, 2) non-reversible transactions, and 3) cost of mediation (transaction costs), for example. The need arose to make electronic payments using cryptography instead of a trusted third party (Nakamoto, 2008). Under this approach, the concept of cryptocurrency became a buzzword both in industry and academia and, in fact, for the general public (Zheng, Xie, Dai, Chen, and Wang, 2017).

Media disseminated concepts such as blockchain and cryptocurrency, relating them to Bitcoin's definition despite their conceptual differences. We can link Bitcoin from its use as a means of payment in the deep web market place Silkroad to be part of the Financial Action Task Force (FATF) in its report called Virtual Currencies Key Definitions and Potential AML / CT Risks. From a general perspective, this financial innovation applicable to different areas of knowledge is impressive for a simple reason: Bitcoin's underlying structure is fascinating (Griffin, 2014).

Consequently, Bitcoin's relevance lies in the potential existing on the effects on payment systems and possibly monetary systems. Likewise, "Bitcoin can be understood as the first widely adopted mechanism to provide absolute scarcity of a money supply" (Böhme, Christin, Edelman, and Moore, 2015, p. 214-215). These authors have detected the risks or possible risks that involve this virtual currency. One of them is relevant in studying financial bubbles because they indicated that Bitcoin retainers face market risk derived from fluctuations in the exchange rate between this cryptocurrency and other currencies. They mentioned that a "user might dismiss the short-term price spikes before mid-2013 as part of the price of using a new currency. But the sharp movements from late 2013 through 2015 would be a source of concern, both for users considering Bitcoin for transactions and for those using it as a store of value" (Böhme et al., 2015, p. 226).

In this sense, Glaser, Zimmermann, Haferkorn, Weber, and Siering (2014) pointed out that in January 2011, each Bitcoin was worth $0.3, while in November 2013, the price increased dramatically to $1,300. Consequently, financial market regulators wondered about the usefulness of this cryptocurrency, referring to the fact that users were not using Bitcoin as a medium of exchange but as a speculative financial asset. Similarly, Swartz (2014) exposes that Bitcoin's demand is increasing based on the public; however, its uses have remained constant. Consequently, according to some exchanges, the acquisition of this cryptocurrency shows its purchases as a speculative tool.

Bitcoin's peculiar history did not stop with the previous explosive price behavior. However, as of December 12, 2017, this virtual currency reached a maximum daily closing price of $18,674.48, according to information extracted from Bloomberg. Thus, concerning the $0.3 value in 2011, this increase meant a 6,224,726.67 percent change, while for a $0.05 price per Bitcoin on July 22, 2010, it represents a 37,348,860 percent increase. It is worth mentioning that their collapses are as distinctive as their rises.

2.3 A STATE-OF-THE-ART APPROACH

2.3.1 BIBLIOMETRIC ANALYSIS

Based on the work of Zhou, Chen, and Huang (2019) where they carried out a scientometric analysis of the financial bubbles of a total of 1,048 articles downloaded from Web of Science (WoS) from the period 1994 to October 26, 2017, with the “financial bubble” research topic. Hence, we carried out a bibliometric analysis using the *bibliometrix* R package, where the study period was extended from 1972 to April 18, 2020, to visualize the largest number of Scopus articles available. For 2,494 articles, the search words were the following: *financial bubble*, *asset price bubble*, *stock market bubble*, or *bubble*. The selection of the subject area was *Economics, Econometrics, and Finance*.

In Fig. 2 and Table 1, we observe the evolution of articles published since 1972 with an annual percentage growth rate of 9.050773. It is worth mentioning that there were no publications in this database during the years 1973, 1974, 1975, 1976, 1978, 1981, and 1982. Although we were only in the fourth month of 2020, the growth in publications of asset price bubbles has been increasing.

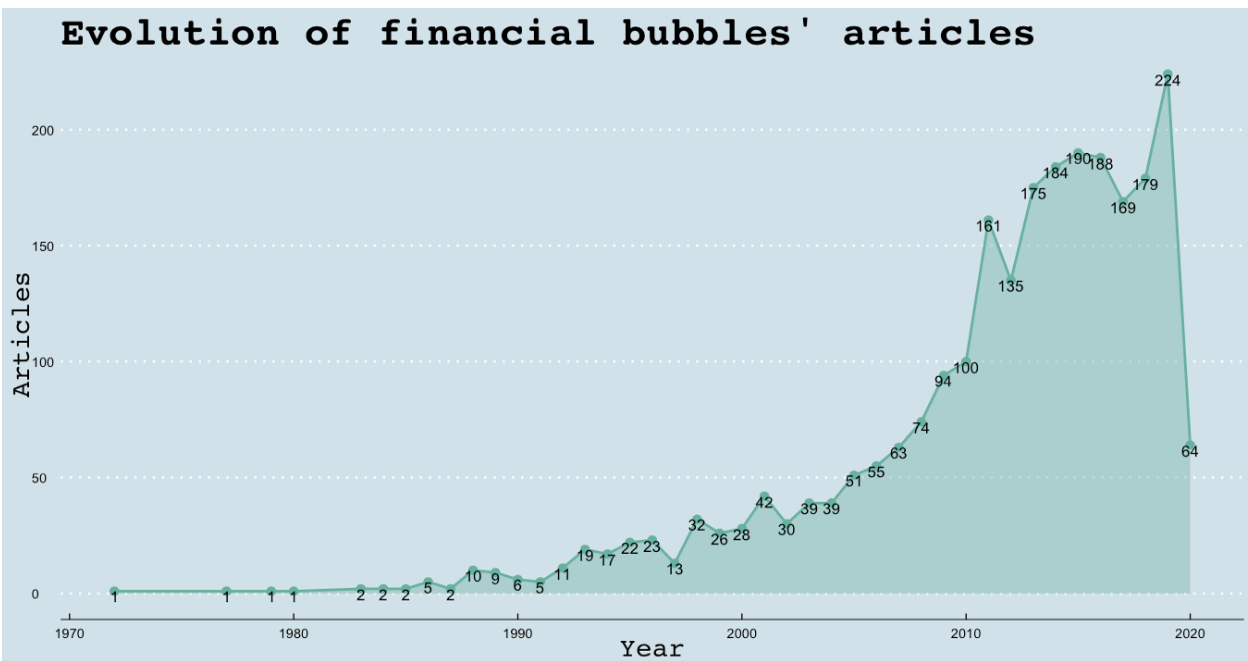

Fig. 2 Number of articles per year.

Table 1. Number of articles per year.

Year	1972	1977	1979	1980	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Articles	1	1	1	1	2	2	2	5	2	10	9	6	5	11

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Articles	19	17	22	23	13	32	26	28	42	30	39	39	51	55

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Articles	63	74	94	100	161	135	175	184	190	188	169	179	224	64

In Table 2, we indicate the top ten of the most productive authors based on the number of articles and articles fractionalized. We omitted the first line of results due to the accumulation of anonymous articles denoted as *NA NA*, where Prof. Didier Sornette stands out in this category. Table 3 shows the first ten articles with the highest number of total citations (TC) and the cited manuscripts by the yearly average number of times (TCperYear).

Table 2. Top ten most productive authors.

Number	Authors	Articles	Authors	Articles Fractionalized
1	Sornette Didier	28	Sornette Didier	11.4
2	Irwin Scott H.	13	Jarrow Robert	8.67
3	Protter Phillip	13	Mcmillan David G.	7.33
4	Su Chi-Wei	13	Irwin Scott H.	6.17
5	Jarrow Robert	12	Westerhoff Frank	5.83
6	Shi Shuping	11	Huang MeiChi	5.5
7	Hommes Cars	10	Protter Phillip	5.33
8	Phillips P.C.B.	10	Tsai I.C.	5.33
9	Westerhoff Frank	10	Engsted Tom	5.17
10	Yu Jun	10	Shi Shuping	5.08

Table 3. Top ten most cited manuscripts.

Number	Authors	Manuscripts	TC	TCperYear
1	De Long, J. Bradford, et al.	Positive feedback investment strategies and destabilizing rational speculation.	1,007	32.5
2	Scheinkman, Jose A., and Wei Xiong.	Overconfidence and speculative bubbles.	686	38.1
3	Boyer, Robert.	Is a finance-led growth regime a viable alternative to Fordism? A preliminary analysis.	467	22.2
4	Engel, Charles.	The forward discount anomaly and the risk premium: A survey of recent evidence.	450	18
5	Chen, Joseph, Harrison Hong, and Jeremy C. Stein.	Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices.	415	20.8
6	Abreu, Dilip, and Markus K. Brunnermeier.	Bubbles and crashes.	393	21.8
7	Brown, John Seely, and Paul Duguid.	Balancing act: How to capture knowledge without killing it.	345	16.4
8	Efendi, Jap, Anup Srivastava, and Edward P. Swanson.	Why do corporate managers misstate financial statements? The role of option compensation and other factors.	343	24.5
9	Phillips, Peter CB, Yangru Wu, and Jun Yu.	Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?	333	33.3
10	Lux, Thomas.	The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions.	329	14.3

In Fig. 3, we show the corresponding authors' countries with the highest number of articles published, the USA tops the list, and for the second place, which is the United Kingdom, it has 257.33 percent more articles. In Table 4, we displayed the most relevant sources, and in Table 5, the most relevant keywords. On the other hand, an interesting fact that calculates the authors' dominance ranking, as proposed by Kumar and Kumar (2008), where Jarrow Robert and Allen Franklin dominate their research for a $k = 10$. Likewise, for Lotka's law coefficients for scientific productivity, the estimated Beta coefficient is 2.435754 with a goodness of fit equal to 0.8773638, and the Kolmogorov-Smirnoff two-sample test provides a p-value of 0.1813004 that means there is not a significant difference between the observed and the theoretical Lotka distributions.

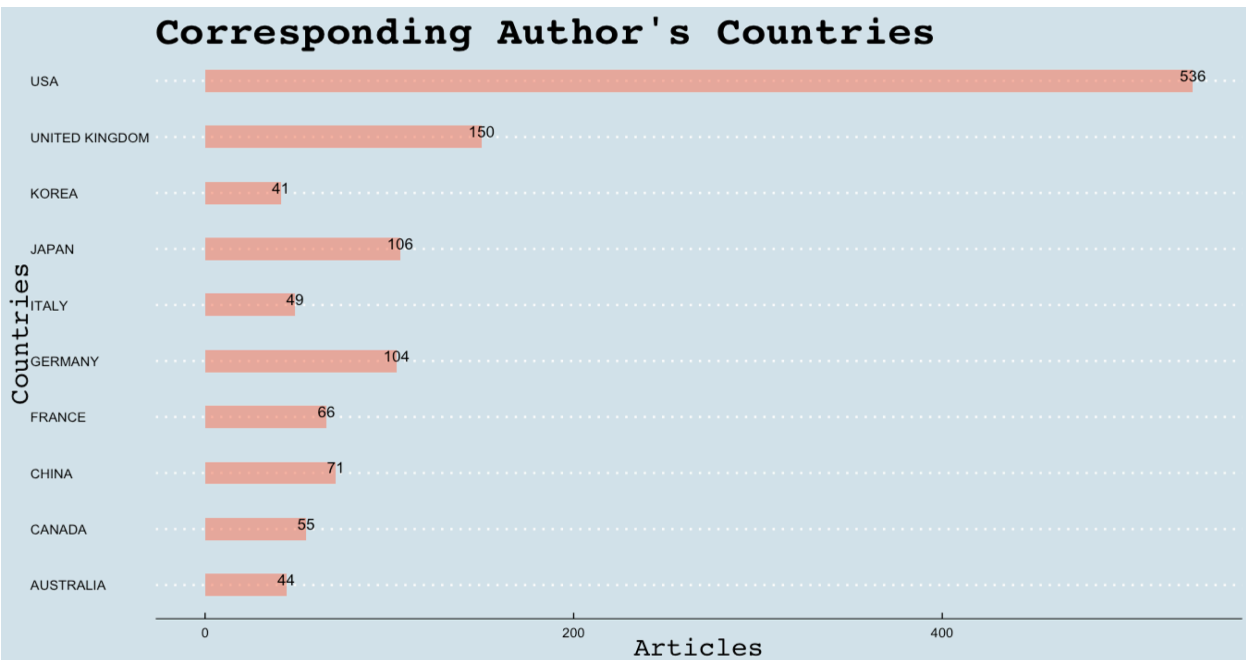

Fig. 3 Corresponding author's countries.

Table 4. Top ten most relevant sources.

Number	Sources	Articles
1	Economics Letters	48
2	Journal of Economic Dynamics and Control	48
3	Economic Modelling	45
4	Journal of Economic Behavior and Organization	36
5	Applied Economics	35
6	Applied Economics Letters	34
7	Journal of Real Estate Finance and Economics	33
8	Quarterly Review of Economics and Finance	30
9	Journal of Banking and Finance	28
10	Journal of Monetary Economics	28

Table 5. Top ten most relevant keywords.

Number	Author Keywords (DE)	Articles
1	Bubbles	236
2	Bubble	101
3	Monetary policy	94
4	Financial crisis	73
5	Speculative bubbles	52
6	Rational bubbles	50
7	House prices	49
8	Housing bubble	48
9	Cointegration	40
10	Asset price bubbles	38

Finally, in Fig. 4, we create a co-citation network analysis for fifty references, displaying the intellectual picture in financial bubble researchers. Simultaneously, in Fig. 5, we developed a keyword co-occurrences analysis of fifty terms, where we find five clusters. In the first cluster (purple color), we have the following keywords: *stochastic systems, bubbles (in fluids), oil prices, oil trade, crude oil, costs, commerce, energy market, price determination, economics, investments, economic impact,* and *commodity price*. In the red cluster, we have the terms of *Markov chain, econometrics, price dynamics, numerical method, market conditions, stock market, housing market,* and *China*. Then, in the blue cluster, words such as *Europe, North America, Japan, Asia, Eurasia,* and *far east* are contemplated.

On the other hand, in the green cluster, the keywords are *investment, United States, European Union, financial market, economic development, inflation, financial crisis, United Kingdom, monetary policy, macroeconomics, central bank, interest rate, financial system, economic growth,* and *banking*. Lastly, we obtained the orange cluster where it presents a different branch of knowledge; however, it is related to some concepts of the clusters mentioned above. We have keywords such as *article, human, humans, controlled study, nonhuman, procedures, chemistry,* and *particle size* in this group.

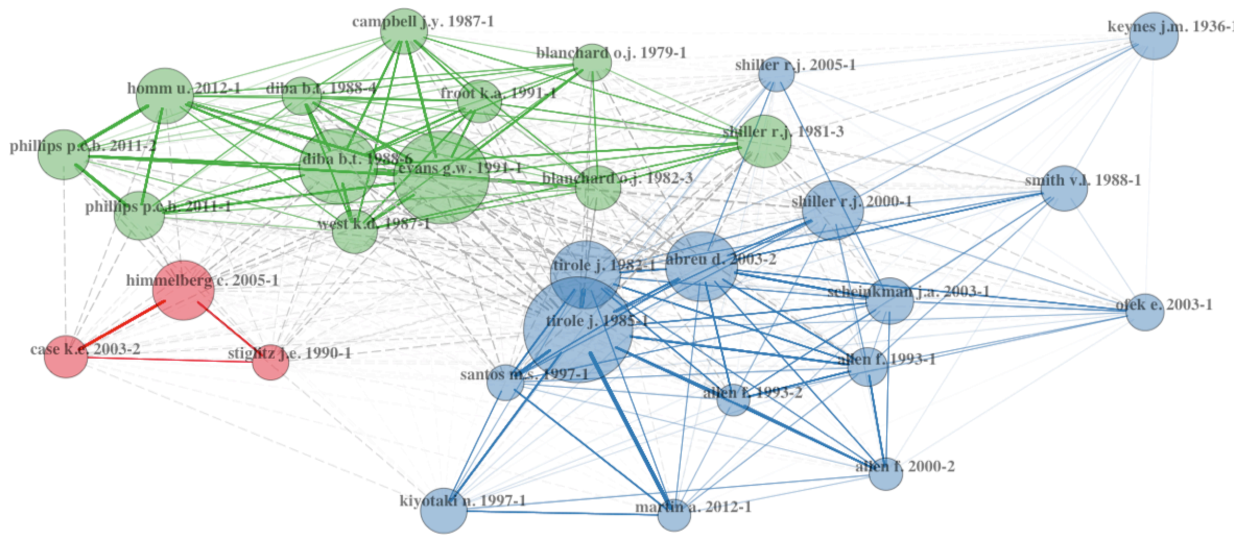

Fig. 4 Co-citation Network of thirty references.

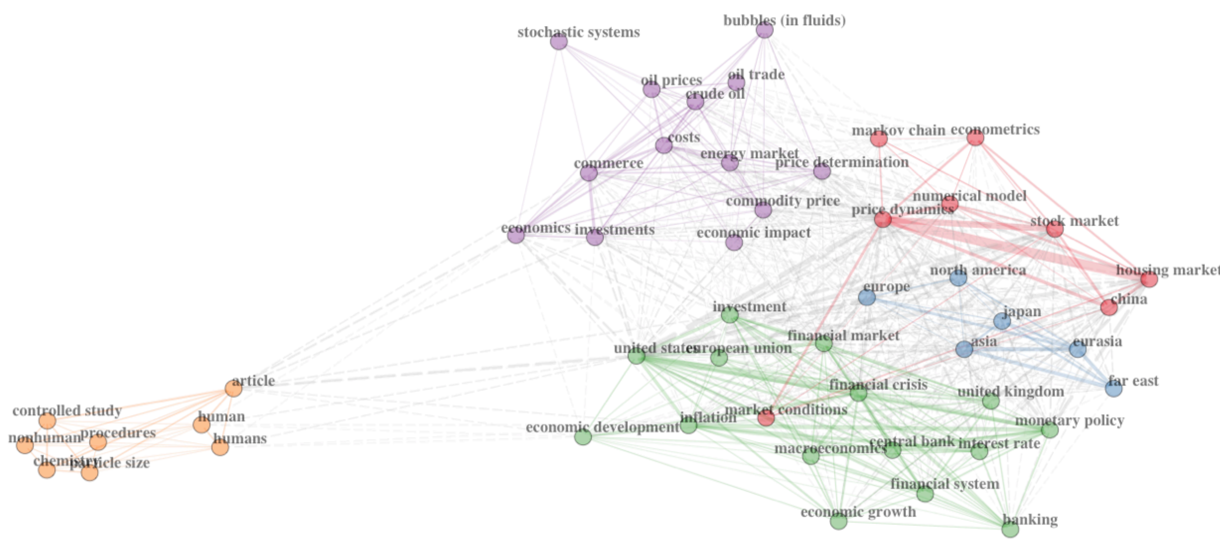

Fig. 5 Keyword Co-ocurrances of fifty terms.

Consequently, from this bibliometric analysis of the financial bubbles, we obtained the general panorama of the investigations carried out in 2,494 articles from 1972 to April 18, 2020. The results in this subject are of ongoing interest, and we can appreciate that more researchers have followed this line of study over time. This type of analysis allows us to highlight the prominent authors who published on asset price bubbles. In our case, it is Professor Didier Sornette and P.C.B. Phillips, as well as to understand that this topic is booming due to the events in early 2020. Moreover, at the time of writing this work, we have witnessed

1. The outbreak of a pandemic (COVID-19),
2. The collapse of different stock market indices, and
3. The price collapse of May's oil futures contracts that turned into negative numbers.
4. The COVID-19 second outbreak.

The previous experiences are a sequence of abnormal events that no one expected to occur. In Sornette's (2017) words, abnormal means "essentially impossible. The fact that they occurred tell us that the market can deviate significantly from the norm" (p. 50).

3. CHAPTER II: POTENTIAL EXUBERANCE IN THE XXI CENTURY

3.1 INTRODUCTION

After the fall of the financial markets in 2008, individuals continue with the same expectations and human desires regarding the acquisition of wealth, greed, and the firm belief that everything done is correct. Nowadays, there is uncertainty if we are still in economic and financial recovery, or perhaps we are in preparation for a magnificent fall of divine proportions, something that has been called in the industry as the everything bubble. In March 2020, the markets' instability got tested; they suffered losses more remarkable than one digit due to the COVID-19 outbreak's uncertainty.

We have not appreciated and learned from past mistakes, especially in the financial markets. Probably there are diverse interests about the real objective of the financial system. Likewise, there exists a possibility that our thinking, our selfish thinking, has inflicted damage to the natural order of the markets. One problem is that we, as human beings, focus on the information from what just occurred yesterday, and we diminish the knowledge importance from one month, one year, or a decade ago. One of the most remarkable socio-economic-political events, the 2008 crisis, has been forgotten like ashes going through the air. It seems that we have put on a shelf those memories like an old book getting dust over time.

The nearest revival of the 2008 crisis was the collapse that accompanied the first pandemic outbreak in 2020. We are talking of the word "accompanied" and not "caused" because causation is not a straightforward job. We cannot explain the natural evolution of financial bubbles as in "real" life with one variable, instead it is an interaction of multiple variables with no linear relations.

We decided to magnify our research on cryptocurrencies (Cerecedo-Hernández, Franco-Ruiz, Contreras-Valdez & Franco-Ruiz, 2019) to a more general context on the potential exuberance in the financial markets within the top fifty of the S&P 500 index regarding their Index weight and starting date January 1, 2000. This chapter was developed by detecting financial bubbles through transitioning from the assignment of a stationary process to a unit root, including a "mildly explosive" process, and returning to a stationary process (Zhang & Wu, 2018). The detection mentioned above derived from the investigations carried out by Phillips, Wu, and Yu (2011) and Phillips, Shi, and Yu (2015), where we developed the Sup Augmented Dickey-Fuller (SADF) and generalized sup Augmented Dickey-Fuller (GSADF) tests, respectively. Likewise, we incorporate the modification proposed by Harvey, Leybourne, Sollis, and Taylor (2016) of the implementation of a wild bootstrap rather than Monte Carlo simulations because the "supremum-based test has a non-pivotal limit distribution under the unit root null, and can be quite severely over-sized, thereby giving rise to spurious indications of explosive behavior" (p. 548).

3.2 THEORETICAL FRAMEWORK

Recently, Phillips et al. (2011, 2015) developed an innovative and persuasive approach to identify bubbles. Phillips et al. (2011), in their seminal paper, defined financial exuberance and introduced a "new econometric methodology based on forward recursive regression tests and mildly explosive regression asymptotics to assess the empirical evidence of exuberant behavior in the Nasdaq stock market index" (p. 202). In order to explain explosive behavior in economic variables, they related their proposal to rational bubble literature. In this context, Diba and Grossman (1988) applied the standard unit root tests to levels and differences of the U.S. Standard and Poor's Composite Stock Price Index data from 1871 to 1986 and rejected the presence of a bubble. So, they mentioned:

> A rational bubble reflects a self-confirming belief that an asset's price depends on a variable (or combination of variables) that is intrinsically irrelevant – that is, not part of market fundamentals – or on truly relevant variables in a way that involves parameters that are not part of market fundamentals. (p. 520).

However, Evans (1991) argued that standard unit root cointegration tests could not distinguish between a stationary process and an occasionally falling bubble model. Thus, these are not adequate tools to identify exuberance. In this context, Phillips et al. (2011) interpreted Greenspan's famous term "irrational exuberance" to signal that the market is overvalued and might be a risk of a financial bubble. One of the possible reasons why explosive behavior does not receive much attention, according to Campbell, Champbell, Campbell, Lo, Lo, and MacKinlay (1997), was that "empirically there is little evidence of explosive behavior" (p. 260). Nevertheless, Evans (1991) stated that exuberance is not a constant; instead, it is a temporary behavior and may likely be a first-order integrated process I(1). Phillips et al. (2011) denoted:

> This article has proposed a new approach to testing for explosive behavior in stock prices that makes use of recursive regression, right-sided unit root tests, and a new method of confidence interval construction for the growth parameter in stock market exuberance… The present econometric methodology shows how the data may be studied as a mildly explosive propagating mechanism. (p. 221-22).

In the same year, 2011; Phillips, and Yu (2011) modified the previously proposed methodology to identify bubble behavior in Phillips et al. (2011). The improvements consist of three main aspects:

1. The initial observation is selected based on an information criterion, allowing sharper identification of the bubble birth date.
2. They developed a method for testing bubble migration; subsequently, they established a new limit theory.
3. The paper studied the subprime crisis.

Phillips, Shi, and Yu (2014) detailed the importance of the null and alternative hypotheses and the regression model specification and give guidelines for implementing right-tailed unit root tests. We can find more tests for explosive behavior in Homm and Breitung (2011), yet they confirmed the SADF outperforms the other tests.

Phillips et al. (2015) developed a new recursive flexible method for long historical time series. They proposed the generalized sup ADF test (GSADF), a version of the SADF, for multiple bubbles instead of just for one explosive behavior. Their findings summarized as follows:

1. SADF and the GSADF tests are consistent in the detection of a single bubble in a certain period.
2. The SADF has a problem when the sample has two bubbles. For the first one, the detection is consistent, whereas, for the second, the estimations are duration-dependent.
3. The GSADF can detect multiple bubbles, and the results hold irrespective of bubble duration.
4. The simulation results corroborate the asymptotic theory, where the GSADF is more reliable than the SADF.

The SADF "procedure uses recursively calculated right-sided unit root test statistics based on an expanding window of observations up to the current data point…" while the GSADF "use a moving window recursion of sup statistics based on a sequence of right-sided unit root tests calculated over flexible windows of varying length taken up to the current data point" (Phillips et al., 2015, p. 1,080).

Later, Lee and Phillips (2016) employed the same econometric methods described above (SADF and GSADF) to investigate the possible impact of bubbles and provide insights into the relationship between bubble-type behavior and financial returns. Phillips (2016) constructed a model of asset market exuberance, and with some modifications, the model can capture the cross-market speculative relation and negative market sentiment.

Moreover, Phillips and Shi (2018) strengthened the GSADF test by exploring different collapse scenarios and proposed an alternative reverse regression for bubble implosion. Likewise, Phillips and Shi (2019) proposed a mechanism for financial market crises and collapses modeling. It is appropriate to capture the abrupt market falls. We will observe the SADF and GSADF application to different markets or cumulative approaches in the next paragraphs.

Pavlidis, Yusupova, Paya, Peel, Martínez-García, Mack, and Grossman (2016) studied the housing market's explosive behavior using real house prices, price-to-income ratios, and price-to-rent ratio in order to analyze the narrative connections with the housing exuberance of the 2008 crisis. They applied the two recursive univariate unit root tests (SADF and GSADF). Moreover, they proposed a novel approach for "a panel setting to exploit the large cross-sectional dimension" (p. 419-420). They found exuberance in periods previous to the global recession. Likewise, Escobari and Jafarinejad (2016) tested for the existence of bubbles in four Real Estate Investment Trust (REIT) using the SADF and GSADF tests. Their results showed statistically significant evidence of speculative behavior in the REIT index, Equity, Mortgage, and Hybrid REITs (p. 224).

Li, Tao, Su, and Lobont (2018) applied the generalized sup Augmented Dickey-Fuller test method (GSADF) to look for explosive bubbles in the Bitcoin (BTC), especially in its price difference between China and the US. They found six bubbles for China and five bubbles for the latter (p. 91-92). Cerecedo-Hernández et al. (2019) extended the previous research for four other cryptocurrencies. We found for Ethereum, Ripple, Bitcoin Cash, and EOS representing the largest market capitalization after Bitcoin with ten, seven, six, and seven explosive bubbles, respectively (p. 726).

Caspi, Katzke, and Gupta (2018) used the GSADF to date-stamp periods of oil price explosivity in the US from January 1876 to January 2014. They established that there should be caution in interpreting results in storable commodities because not necessarily an explosive period has to be a bubble. Moreover, these authors mentioned there exist a possibility that explosivity as a bubble identification could be an "indicative of adjustments from previously managed or manipulated pricing schedules toward a more fundamental level (whichever way defined)" (p. 583-584). On the other hand, Sharma and Escobari (2018) found strong evidence of explosive episodes in three energy indices (crude oil, heating oil, and natural gas) and five energy spot prices (West Texas Intermediate, Brent, heating oil, natural gas, and jet fuel) obtaining the critical values by Monte Carlo simulations (p. 419).

Hu and Oxley (2018) took two historical crashes, the well-known South Sea Company and the Mississippi Company, as well as six underresearched 18th-century financial series. They applied the right-tailed unit root test (GSADF) and considered non-stationary volatility where critical values originate in a wild bootstrap simulation developed by Harvey et al. (2016).

Chang, Gil-Alana, Aye, Gupta, and Ranjbar (2016) explored the GSADF test for the BRICS (Brazil, Russia, India, China, and South Africa) stock markets 1990-2013 period. Furthermore, Hu and Oxley (2017) used the generalized sup ADF (GSADF) to investigate the exchange rate bubbles in some G10, Asian, and BRICS countries from March 1991 to December 2014. They pointed that "explosiveness in the asset price does not, on its own, imply the existence of rational bubbles, where it is necessary to consider the role player by economic fundamentals in asset prices" (p. 439).

Escobari, Garcia, and Mellado (2017) searched for exuberance in Latin American equity markets, where bubbles appear to begin earlier and stay for a more extended period than the S&P 500 index. They proposed a similar recursive procedure based on Phillips-Perron. They concluded that the Augmented Dickey-Fuller-based and the Phillips-Perron-based tests coincide 92.9% of the times in the analyzed sample.

Long, Li, and Li (2016) determined that gold's price reacts faster to political and economic uncertainties than other commodities and mentioned that, in general, the GSADF method locates explosiveness more accurately than the SADF. Furthermore, once again, they confirmed gold is a haven for investors where its price emerged rapidly “soon after the beginning of the subprime mortgage crisis as panic investors transferred assets from property market to the gold market and the gold price reacted to economic turmoil more rapidly than prices of other commodities” (p. 1,162).

More recently, Kurozumi (2020) compared the ADF and CUSUM type detectors and proves that the local asymptotic theory is more helpful in understanding the properties of these tests. The former is better to detect middle to late breaks, while the latter is suitable for detecting exuberance in early and short-range. Monschang and Wilfling (2020) showed Monte Carlo simulations exhibit substantial size distortions and that the backward SADF (BSADF) test used to date-stamp exuberance proposed in Phillips et al. (2011) outperforms variants such as the sign-based test statistic of Harvey, Leybourne, and Zu (2020). Nevertheless, we should be careful because the BSADF tends to date-stamp non-existing bubble-type behavior, and revealed data frequency is a sensible choice to the practitioner of these tests.

Therefore, the evidence shows the SADF and the GSADF are tests with increasing popularity in different scientific fields. Exuberance identification in asset prices or in any other academic subjects is a complex assignment. Just to give an example, Kräussl, Lehnert, and Martelin (2016) investigated after the record-breaking prices in the art market is there was a speculative bubble. To answer this question, they tested Phillips et al. (2011) modeling approach and identified two historical speculative bubbles and explosive behavior in today’s fine art market segments (p. 99).

3.3 TESTING FOR EXPLOSIVE BEHAVIOR

3.3.1 AUTOREGRESSIVE PROCESS

First of all, we have to make a review of an Autoregressive Process. The notation follows Hamilton (1994).

A first-order autoregression AR (1) satisfies the following difference equation:

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t \tag{1}$$

where, ε_t has mean zero and variance σ^2, and they are not correlated

$$E(\varepsilon_t) = 0 \tag{2}$$

$$E(\varepsilon_t^2) = 0 \tag{3}$$

$$E(\varepsilon_t, \varepsilon_\tau) = 0 \tag{4}$$

For an AR (2),

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \tag{5}$$

or, in lag operator notation,

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = c + \varepsilon_t \tag{6}$$

Equation (5) is stable provided that the roots of

$$(1 - \phi_1 z - \phi_2 z^2) = 0 \tag{7}$$

lie outside the unit circle, so if this condition is satisfied, the AR (2) process is covariance-stationary.

Following (5), for an AR(p),

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \tag{8}$$

with roots,

$$\left(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p\right) = 0 \tag{9}$$

Now, following Pavlidis et al. (2016) and Caspi (2013) in order to have a usual notation of the SADF and GSADF tests, we made some arranges to (8),

$$\Delta Y_t = \alpha_{r1,r2} + \beta_{r1,r2} Y_{t-1} + \sum_{j=1}^{k} \psi^j{}_{r1,r2} \Delta Y_{t-j} + \varepsilon_t \tag{10}$$

where Y_t denotes a generic times series so that it can be the price or, for example, a ratio, ΔY_{t-j} for $j = 1, \dots, k$ are the differenced lags of the time series and $\varepsilon_t \sim N(0, \sigma^2_{r1,r2})$ is the error term. Likewise, $r1$ and $r2$ denote fractions of the total sample size, k is the maximum number of lags included in the specification, and $\alpha_{r1,r2}$, $\beta_{r1,r2}$, and $\psi^j{}_{r1,r2}$ are regression coefficients. Finally, r_w is the fractional window size of the regression, defined by $r_w = r2 - r1$, and $r0$ is an initial fixed window.

3.3.2 SADF TEST

The change from a random walk with a drift to a mildly explosive behavior indicates the origin of exuberance under the assumption that fundamentals belong to a class of first-order integrated process I(1). Equation (10) is similar to the Augmented Dickey-Fuller test, but now we are considering an expanding window and a right-tailed version of the standard unit root test. In consequence, we are interested in testing the null hypothesis, $H_0 = \beta_{r1,r2} = 0$, while the alternative hypothesis $H_1 = \beta_{r1,r2} > 0$ tells us it is a mildly explosive behavior.

If we set $r1$ and $r2$ as fixed to the first and last observations respectively, such that

$$r_w = r0 = 1 \tag{11}$$

We obtained a right-tailed version of the standard Augmented Dickey-Fuller unit root test. Nevertheless, the critical values for testing the null hypothesis differ from the usual one because we need the right tail of the test. Fig. 6 is a graphic example of the ADF procedure.

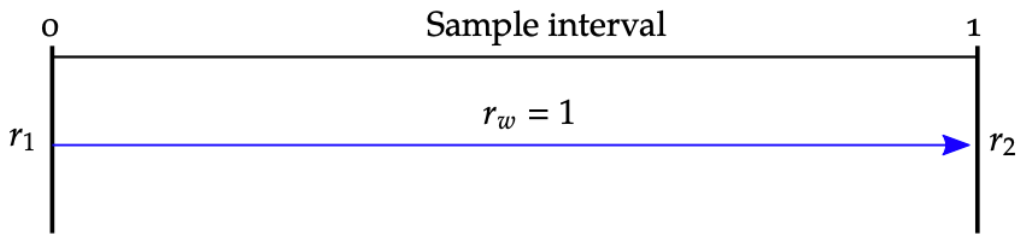

Fig. 6 Illustration of the Augmented Dickey-Fuller procedure extracted from Caspi (2013).

On the other hand, for the SADF test, we established, $r1$, as the starting point of the estimation window. Second, we determined the endpoint, $r2$, according to a minimum window size choice $r0$. Finally, we estimated the regression recursively and kept increasing the window size $r2\epsilon[r0,1]$. The previous process repeats itself one observation at a time. We will get for each estimation an ADF statistic denoted as ADF_{r2}, until it covers the whole sample.

Next, we defined the SADF's test statistic as the supremum value of ADF_{r2}, expressed as follows:

$$SADF(r0) = sup_{r2\epsilon[r0,1]}\{ADF_{r2}\} \quad (12)$$

Equation (12) tells us the SADF will find the largest ADF statistic through all the expanding window. Hence, if the right-tailed test is sufficiently large, we will get at least one explosive behavior. Fig. 7 is a graphic example of the SADF procedure.

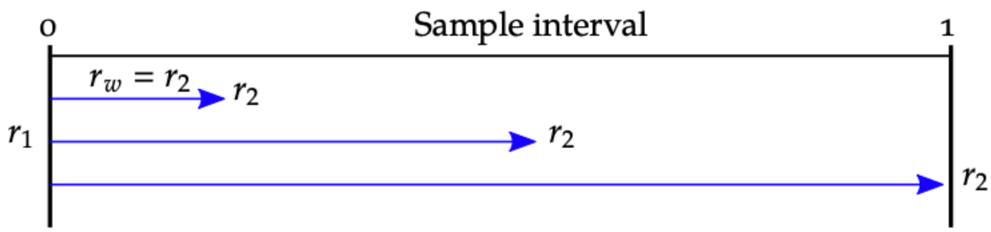

Fig. 7 Illustration of the SADF procedure extracted from Caspi (2013).

3.3.3 GSADF TEST

For the multiple periodically collapsing bubbles, GSADF, the test takes a more flexible window where $r1$ can vary. The GSADF allows $r1$ and $r2$ to change, gaining power against the SADF. We defined the GSADF's test statistic as follows, while Fig. 8 is a graphic example of the GSADF procedure. Likewise, Phillips et al. (2015) recommended a dating strategy that includes a backward SADF statistic (BSADF) to improve accuracy.

$$GSADF(r0) = sup_{r2\epsilon[r0,1],r1\epsilon[0,r2-r0]}\{ADF_{r1}^{r2}\} \quad (13)$$

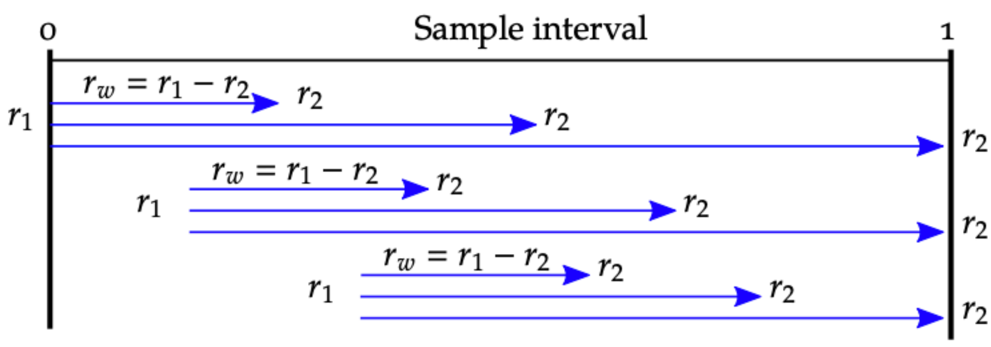

Fig. 8 Illustration of the GSADF procedure extracted from Caspi (2013).

The BSADF statistic relates to the GSADF statistic,

$$GSADF(r0) = sup_{r2\epsilon[r0,1]}\{BSADF_{r2}(r0)\} \quad (14)$$

3.4 POTENTIAL EXUBERANCE IN THE S&P 500 INDEX

We carried out an analysis of the fifty shares with the most significant weight that constitute the S&P 500 index of May 14, 2020. We studied the potential exuberance with the SADF and GSADF methodology since the beginning of the 21st-century. With this S&P 500 index sample, we consider a little more than 55 percent of its constituents (55.17 percent). One of the primary motivations for carrying out this analysis arises from the research where we detected explosive behaviors in Ethereum, Ripple, Bitcoin Cash, and EOS (Cerecedo-Hernández et al., 2019). However, that the combined capitalization of cryptocurrencies based on the study by Mossavar-Rahmani, Nelson, Weir, Minovi, Ubide, Asl, Dibo, and Rich (2018) is less than one percent of world GDP, the innovation that represents the internal structure of cryptocurrencies deserved special investigation and scrutiny.

Therefore, the study of some stocks of the S&P 500 index is transcendental because it is one of the most representative indices of the real situation of the United States market, which according to the website of the S&P Dow Jones Indices, a division of S&P Global, stipulates that the S&P 500 "is widely regarded as the best single gauge of large-cap US equities. There is over USD 9.9 trillion indexed or benchmarked to the index, with indexed assets comprising approximately USD 3.4 trillion of this total. The index includes 500 leading companies and covers approximately 80% of available market capitalization."[3] It has a composition of 505 companies between different sectors, according to the same official page mentioned above. We can observe the sector breakdown in Fig. 9.

[3] https://us.spindices.com/indices/equity/sp-500

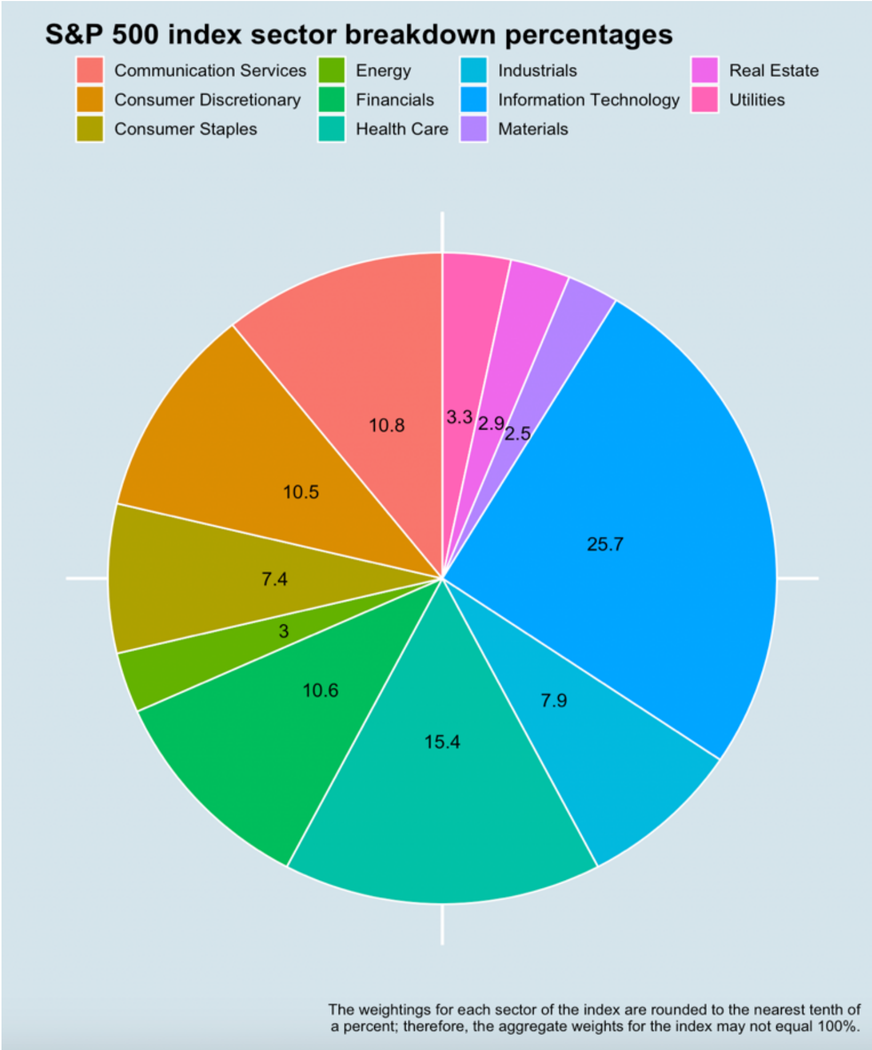

Fig. 9 S&P 500 index sector breakdown percentages.

The fifty shares studied were the following. We extracted the daily adjusted prices from Yahoo Finance. The next constituents of the S&P 500 index are in order by weight.[4]

1. Microsoft Corporation / MSFT
2. Apple Inc. / AAPL
3. Amazon.com Inc. / AMZN
4. Facebook Inc. Class A / FB
5. Alphabet Inc. Class A / GOOGL
6. Alphabet Inc. Class C / GOOG
7. Johnson & Johnson / JNJ
8. Berkshire Hathaway Inc. Class B / BRK-B
9. Visa Inc. Class A / V
10. Procter & Gamble Company / PG
11. UnitedHealth Group Incorporated / UNH

[4] https://www.slickcharts.com/sp500

12. JPMorgan Chase & Co. / JPM
13. Intel Corporation / INTC
14. Home Depot Inc. / HD
15. Mastercard Incorporated Class A / MA
16. Verizon Communications Inc. / VZ
17. Pfizer Inc. / PFE
18. AT&T Inc. / T
19. Merck & Co. Inc. / MRK
20. NVIDIA Corporation / NVDA
21. Netflix Inc. / NFLX
22. Walt Disney Company / DIS
23. Cisco Systems Inc. / CSCO
24. PepsiCo Inc. / PEP
25. Exxon Mobil Corporation / XOM
26. Bank of America Corp / BAC
27. Walmart Inc. / WMT
28. Adobe Inc. / ADBE
29. Chevron Corporation / CVX
30. PayPal Holdings Inc. / PYPL
31. Coca-Cola Company / KO
32. Comcast Corporation Class A / CMCSA
33. Abbott Laboratories / ABT
34. AbbVie Inc. / ABBV
35. Bristol-Myers Squibb Company / BMY
36. salesforce.com inc. / CRM
37. Amgen Inc. / AMGN
38. Thermo Fisher Scientific Inc. / TMO
39. Eli Lilly and Company / LLY
40. Costco Wholesale Corporation / COST
41. McDonald's Corporation / MCD
42. Medtronic Plc / MDT
43. Oracle Corporation / ORCL
44. Accenture Plc Class A / ACN
45. NextEra Energy Inc. / NEE
46. NIKE Inc. Class B / NKE
47. Union Pacific Corporation / UNP
48. Broadcom Inc. / AVGO
49. Philip Morris International Inc. / PM
50. International Business Machines Corporation / IBM

From this database, we obtained the descriptive statistics of the daily adjusted prices and their correlations. Table 6, and Annex I show the summary.

Table 6. S&P 500 index daily adjusted prices sample descriptive statistics.

#	Symbol	Sample	Mean	Standard deviation	Median	Minimum	Maximum	Range	Skewness	Kurtosis
1	AAPL	5,113	59.24	69.10	26.59	0.81	326.32	325.51	1.39	1.40
2	ABBV	1,844	58.37	18.91	53.14	24.65	108.30	83.66	0.39	-0.89
3	ABT	5,113	26.73	20.45	17.47	5.62	98.00	92.38	1.38	1.14
4	ACN	4,724	61.85	51.60	37.29	8.87	214.95	206.08	1.15	0.23
5	ADBE	5,113	70.40	78.96	35.29	8.32	383.28	374.96	2.00	2.97
6	AMGN	5,113	84.27	52.34	55.11	25.10	240.80	215.70	1.07	-0.19
7	AMZN	5,113	387.38	554.09	125.83	5.97	2,410.22	2,404.25	1.79	1.98
8	AVGO	2,701	114.34	94.87	87.63	11.83	319.27	307.44	0.56	-1.15
9	BAC	5,113	21.02	9.79	18.38	2.80	42.69	39.89	0.41	-0.93
10	BMY	5,113	29.89	16.82	24.49	10.27	68.82	58.54	0.64	-1.09
11	BRK-B	5,113	98.29	53.35	78.91	27.40	230.20	202.80	0.91	-0.41
12	CMCSA	5,113	16.76	11.52	10.41	4.54	47.19	42.65	0.97	-0.47
13	COST	5,113	85.17	71.04	50.02	19.82	323.34	303.52	1.43	1.30
14	CRM	3,991	52.05	47.53	36.56	2.40	193.36	190.96	1.06	0.12
15	CSCO	5,113	22.51	11.49	18.33	6.59	61.40	54.80	1.34	0.75
16	CVX	5,113	61.39	32.42	56.78	15.06	122.60	107.54	0.16	-1.29
17	DIS	5,113	50.75	37.32	29.68	10.83	150.74	139.90	0.89	-0.65
18	FB	1,999	112.92	58.20	114.60	17.73	223.23	205.50	-0.05	-1.27
19	GOOG	3,951	507.04	356.37	330.75	49.82	1,526.69	1,476.87	0.91	-0.40
20	GOOGL	3,951	512.54	359.80	332.32	50.06	1,524.87	1,474.81	0.88	-0.48
21	HD	5,113	64.73	59.02	30.79	13.58	245.38	231.79	1.34	0.55
22	IBM	5,113	98.41	35.81	91.49	35.96	163.32	127.36	0.10	-1.58
23	INTC	5,113	23.53	12.18	18.55	8.48	67.75	59.27	1.28	0.78
24	JNJ	5,113	63.98	34.83	46.69	20.18	155.51	135.33	0.94	-0.52
25	JPM	5,113	44.15	28.29	32.29	9.38	138.75	129.37	1.47	1.14
26	KO	5,113	24.10	13.03	18.65	8.66	59.61	50.94	0.58	-0.92
27	LLY	5,113	48.84	26.56	38.38	18.26	162.17	143.91	1.57	1.98
28	MA	3,506	79.72	78.56	51.74	3.29	344.03	340.74	1.33	0.90
29	MCD	5,113	66.23	53.42	47.69	7.66	217.27	209.61	1.09	0.29
30	MDT	5,113	48.60	21.59	38.57	19.02	120.55	101.53	1.20	0.39
31	MRK	5,113	37.25	17.14	31.31	14.17	91.29	77.11	1.13	0.62
32	MSFT	5,113	38.04	33.81	22.44	11.70	188.19	176.49	2.16	4.10
33	NEE	5,113	63.68	56.98	40.75	7.63	282.22	274.59	1.50	1.65
34	NFLX	4,515	72.52	109.88	14.52	0.37	439.17	438.80	1.72	1.68
35	NKE	5,113	25.49	26.47	12.61	1.00	104.29	103.29	1.09	0.02

36	NVDA	5,113	44.37	69.30	14.37	2.26	314.51	312.25	2.06	2.90
37	ORCL	5,113	26.15	13.53	24.70	6.34	59.10	52.75	0.46	-0.93
38	PEP	5,113	59.21	30.60	48.71	18.66	146.00	127.34	0.90	-0.28
39	PFE	5,113	20.86	8.25	18.55	7.67	43.69	36.02	0.80	-0.25
40	PG	5,113	52.08	24.48	46.19	14.32	126.30	111.98	0.85	0.40
41	PM	3,052	60.38	21.55	63.30	19.01	104.99	85.97	-0.18	-0.90
42	PYPL	1,214	69.50	28.70	73.94	30.63	123.91	93.28	0.18	-1.44
43	T	5,113	20.15	7.70	18.48	7.77	38.43	30.66	0.44	-0.91
44	TMO	5,113	84.21	77.76	50.64	11.78	342.71	330.93	1.42	1.16
45	UNH	5,113	74.44	75.19	43.36	5.00	303.99	298.99	1.47	0.92
46	UNP	5,113	50.00	48.78	27.10	3.57	186.09	182.53	1.05	0.00
47	V	3,050	63.99	51.57	49.94	8.48	212.95	204.47	0.95	-0.15
48	VZ	5,113	27.51	13.34	20.92	10.31	60.78	50.46	0.76	-0.69
49	WMT	5,113	53.01	21.77	42.19	29.88	131.75	101.87	1.36	1.16
50	XOM	5,113	51.38	19.06	54.66	18.01	82.04	64.04	-0.31	-1.32

As we can notice, the number of observations depends on the available information; our study's most extensive data set has 5,113 daily adjusted prices from January 3, 2000, to April 29, 2020. The S&P 500 index sample had significant upward changes, as we can analyze from Table 6; thus, the average percentual increase from the minimum to the maximum price is over six thousand. The five stocks with the most remarkable growth are NFLX, AMZN, AAPL, NVDA, and MA.

In Annex I, we calculated the correlation matrix of the daily adjusted prices from July 6, 2015, to April 29, 2020, because the series in this period share the same start date. Interestingly, Bristol-Myers Squibb Company (BMY) with 44 of the 49 companies and Exxon Mobil Corporation (XOM) with 40 of the 49 companies are negatively correlated. On the other hand, AT&T Inc. (T) has in all observations a positive correlation of less than 0.7, and International Business Machines Corporation (IBM) has a positive correlation of less than 0.7 with 47 of 49 companies. We found 1,846 of 2,500 observations from the correlation matrix with a value higher than 0.7, excluding for obvious reasons the main diagonal, meaning there is a strong linear relation in the sample.

We use last month's price for each stock to change the daily data to a monthly data set. We employed the SADF and GSADF tests to identify exuberance and date bubble-type periods with the new data set. We made calculations of the model with 2,000 replications and applied wild bootstrap to obtain the critical values. First, we ran the SADF as an initial method to analyze if we could identify at least one explosive behavior. Despite the rejection of the SADF's null hypothesis, we only consider the possible existence of a bubble-type behavior if we could reject the null hypothesis of the GSADF test.

Nevertheless, in Table 7, we identified the stocks that could no reject the null hypothesis of the SADF as "N", while the stocks that reject the null hypothesis of the SADF but not the GSADF, we denoted them as "N*". We decided on this condition because, as mentioned above, the GSADF outperforms the SADF in detecting explosive behavior. Therefore, for the fifty stocks we analyzed, only for fifteen assets, we can assert the possible presence of exuberance.

We interpret the start date of the bubble-type when the GSADF statistic (blue line) is above the critical value (red line) while we defined the end date when the GSADF statistic (blue line) goes below the critical value (red line) and highlight the date-stamp with "pink" color. We graph the estimated result for Microsoft Corporation according to the GSADF test, with 95% confidence intervals in Fig. 10. For this stock, we identified two possible explosive behavior:

1. From July 31, 2014 to January 30, 2015.
2. Beginning in January 21, 2017 but has not ended yet.

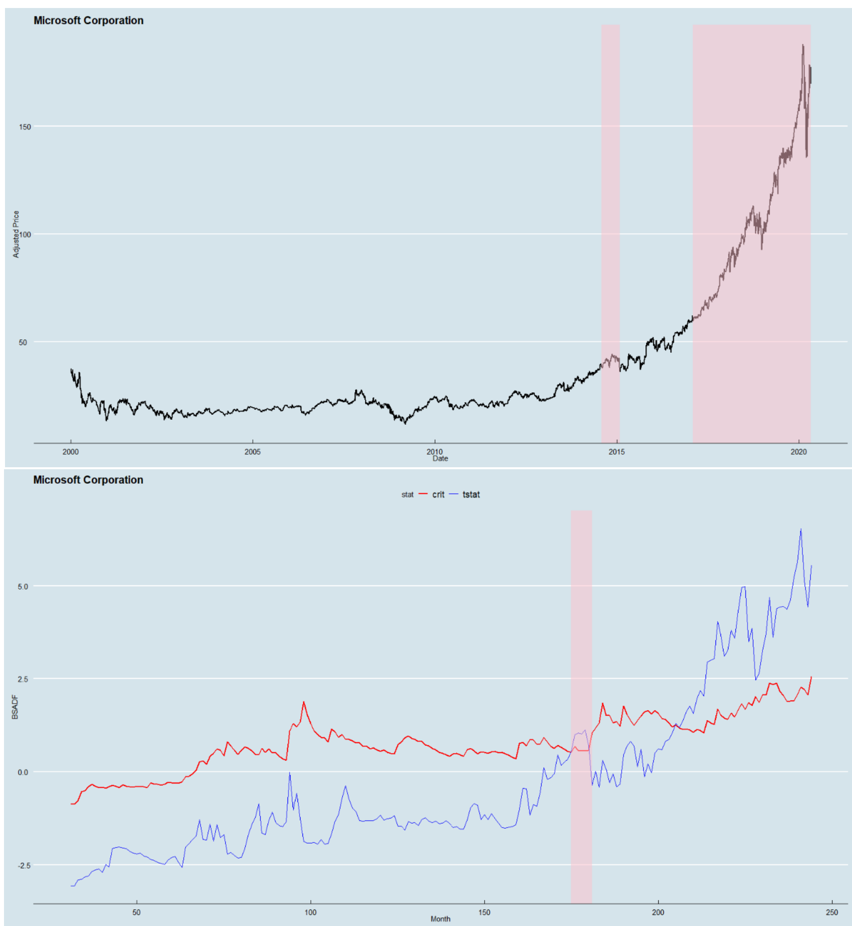

Fig. 10 In the image above, we noticed the adjusted daily price (black line) of Microsoft Corporation since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the for forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Amazon according to the GSADF test, with 95% confidence intervals in Fig. 11. For this stock, we identified ten possible bubbles:

1. From June 30, 2003 to February 27, 2004.
2. From December 31, 2010 to January 31, 2011.
3. From April 29, 2011 to November 30, 2011.
4. From September 28, 2012 to October 31, 2012.
5. From February 28, 2013 to April 30, 2013.
6. From May 31, 2013 to August 30, 2013.
7. From September 30, 2013 to March 31, 2014.
8. From June 30, 2015 to August 31, 2015.
9. From September 30, 2015 to January 29, 2016.
10. From May 31, 2016 to October 31, 2018.

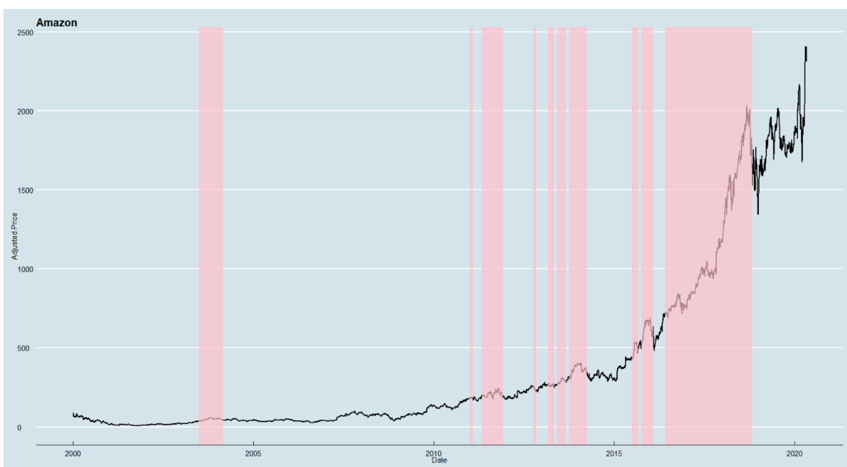

Fig. 11 In the image above, we noticed the adjusted daily price (black line) of Amazon.com Inc. since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Visa Class A according to the GSADF test, with 95% confidence intervals in Fig. 12. For this stock, we identified eight possible bubbles:

1. From February 29, 2012 to May 31, 2012.
2. From July 31, 2012 to April 30, 2014.
3. From May 30, 2014 to August 31, 2015.
4. From October 30, 2015 to January 29, 2016.
5. From March 31, 2016 to June 30, 2016.
6. From August 31, 2016 to November 30, 2016.
7. From February 28, 2017 to December 31, 2018.
8. From January 31, 2019 to March 31, 2020.

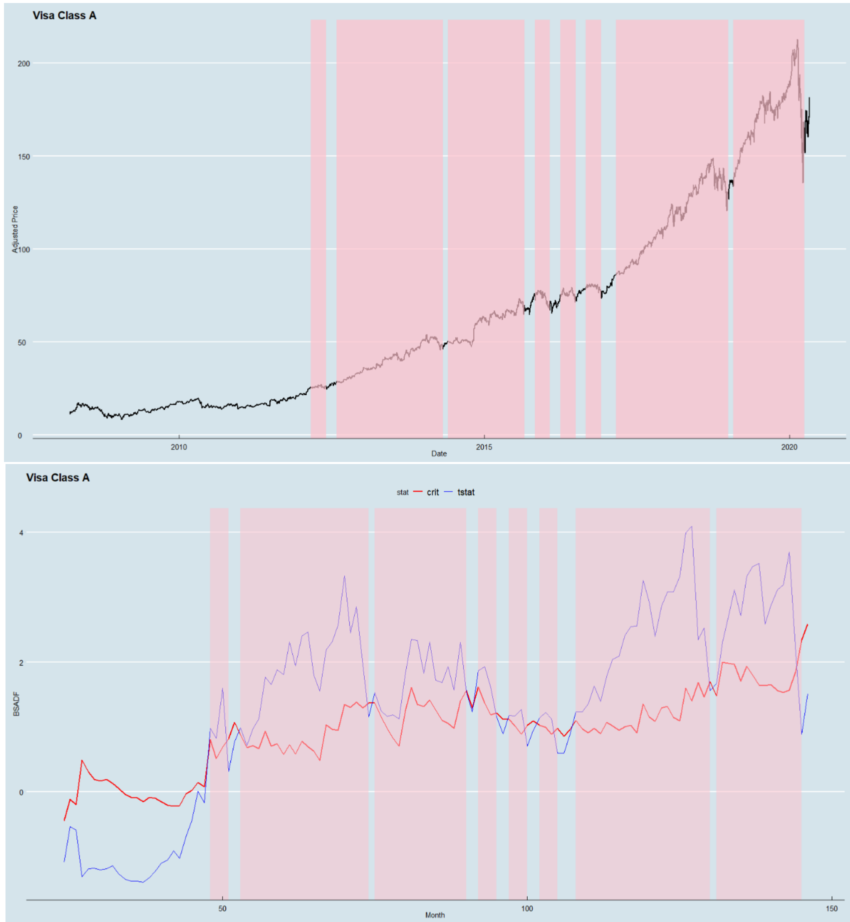

Fig. 12 In the image above, we noticed the adjusted daily price (black line) of Visa Inc. Class A since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for UnitedHealth Group according to the GSADF test, with 95% confidence intervals in Fig. 13. For this stock, we identified six possible bubbles:

1. From February 27, 2004 to April 30, 2004.
2. From November 30, 2004 to March 31, 2006.
3. From June 30, 2011 to July 29, 2011.
4. From March 30, 2012 to April 30, 2012.
5. From March 31, 2014 to April 30, 2014.
6. From October 31, 2014 to February 28, 2019.

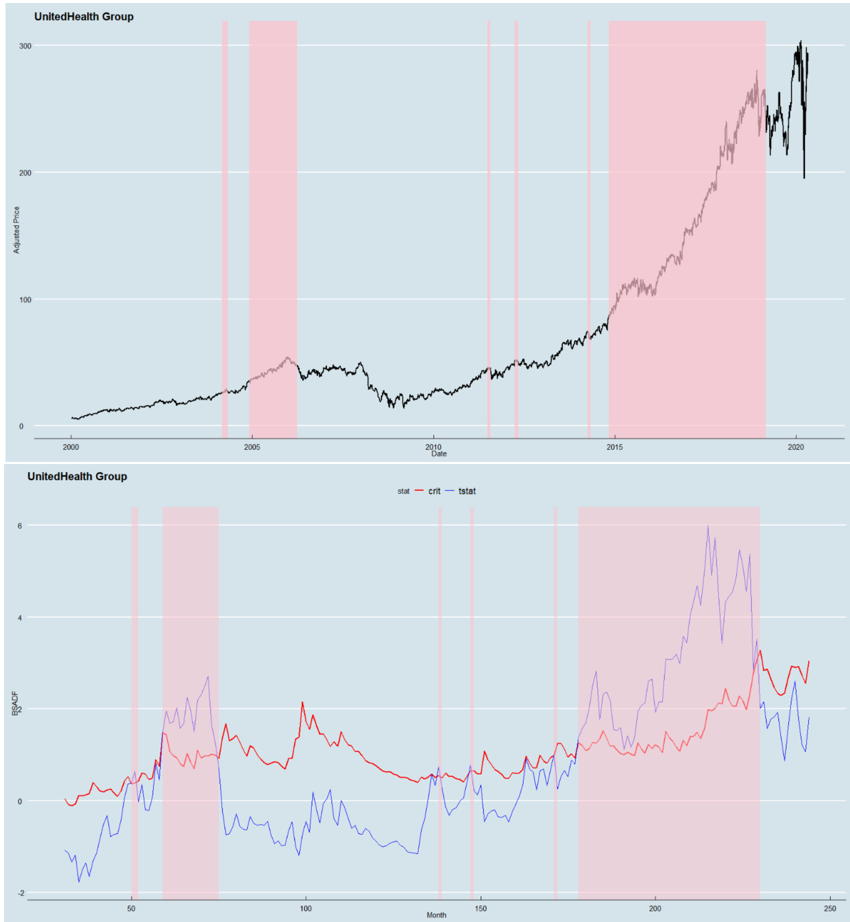

Fig. 13 In the image above, we noticed the adjusted daily price (black line) of UnitedHealth Group Incorporated since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Home Depot according to the GSADF test, with 95% confidence intervals in Fig. 14. For this stock, we identified four possible bubbles:

1. From February 29, 2012 to October 31, 2016.
2. From November 30, 2016 to October 31, 2018.
3. From April 30, 2019 to May 31, 2019.
4. From June 28, 2019 to March 31, 2020.

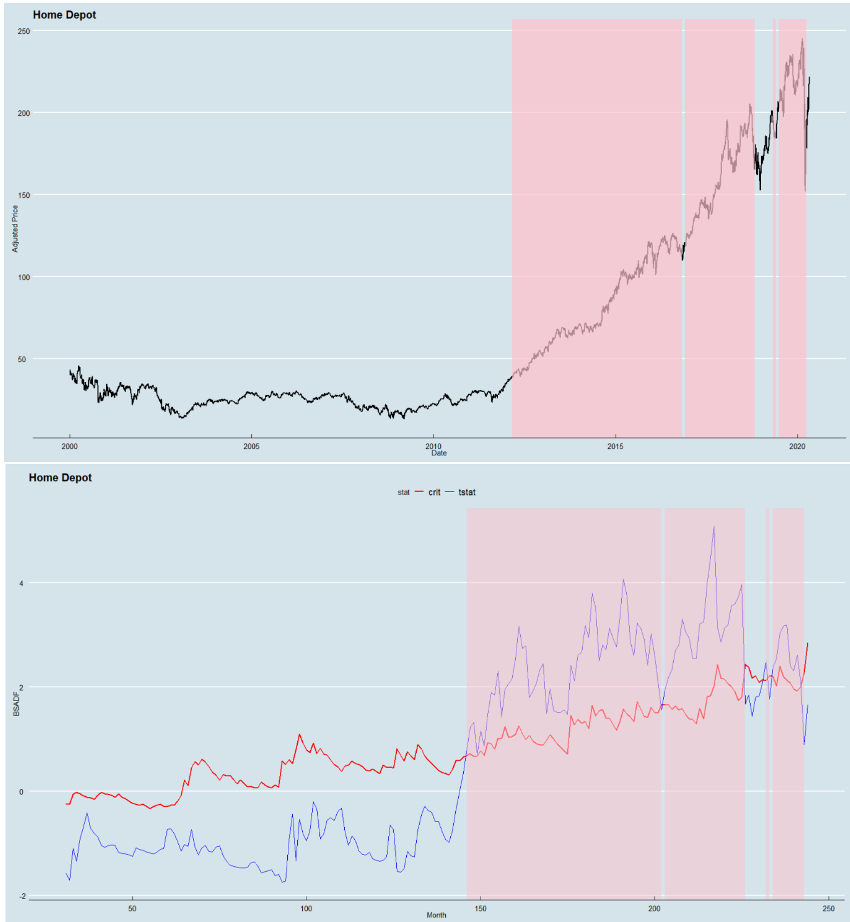

Fig. 14 In the image above, we noticed the adjusted daily price (black line) of Home Depot Inc. since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Mastercard according to the GSADF test, with 95% confidence intervals in Fig. 15. For this stock, we identified seven possible bubbles:

1. From August 31, 2011 to September 30, 2011.
2. From November 30, 2011 to January 31, 2012.
3. From March 30, 2012 to May 31, 2012.
4. From November 30, 2012 to March 31, 2015.
5. From April 30, 2015 to January 29, 2016.
6. From April 28, 2017 to December 31, 2018.
7. From February 28, 2019 to March 31, 2020.

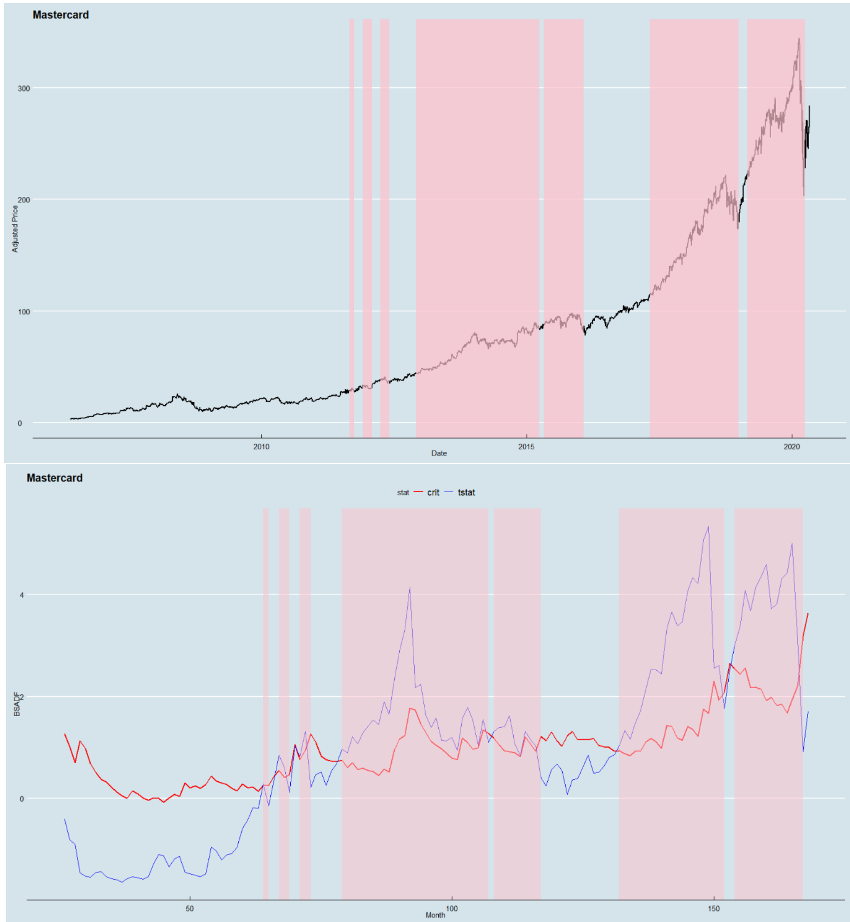

Fig. 15 In the image above, we noticed the adjusted daily price (black line) of Mastercard Incorporated Class A since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for NVIDIA Corporation according to the GSADF test, with 95% confidence intervals in Fig. 16. For this stock, we identified four possible bubbles:

1. From February 28, 2006 to May 31, 2006.
2. From August 31, 2007 to November 30, 2007.
3. From November 30, 2015 to January 29, 2016.
4. From May 31, 2016 to October 31, 2018.

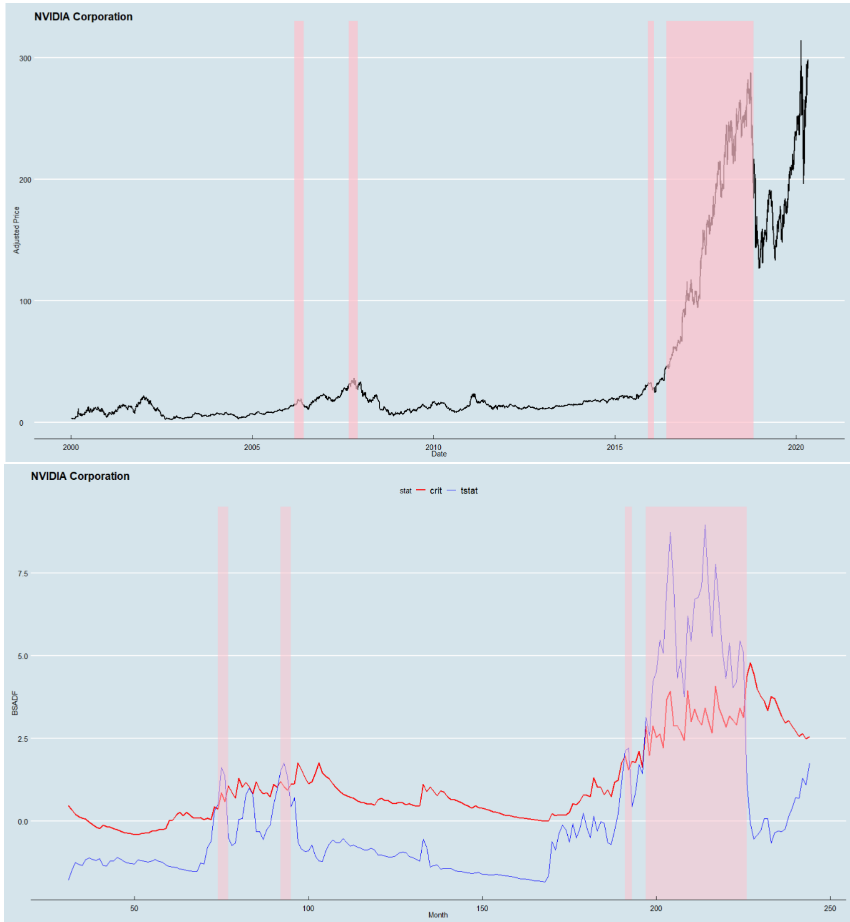

Fig. 16 In the image above, we noticed the adjusted daily price (black line) of NVIDIA Corporation since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Netflix according to the GSADF test, with 95% confidence intervals in Fig. 17. For this stock, we identified nine possible bubbles:

1. From February 26, 2010 to December 31, 2010.
2. From March 31, 2011 to August 31, 2011.
3. From January 31, 2014 to March 31, 2014.
4. From July 31, 2015 to September 30, 2015.
5. From November 30, 2015 to December 31, 2015.
6. From March 31, 2017 to June 30, 2017.
7. From July 31, 2017 to August 31, 2017.

8. From September 29, 2017 to July 31, 2018.
9. From August 31, 2018 to October 31, 2018.

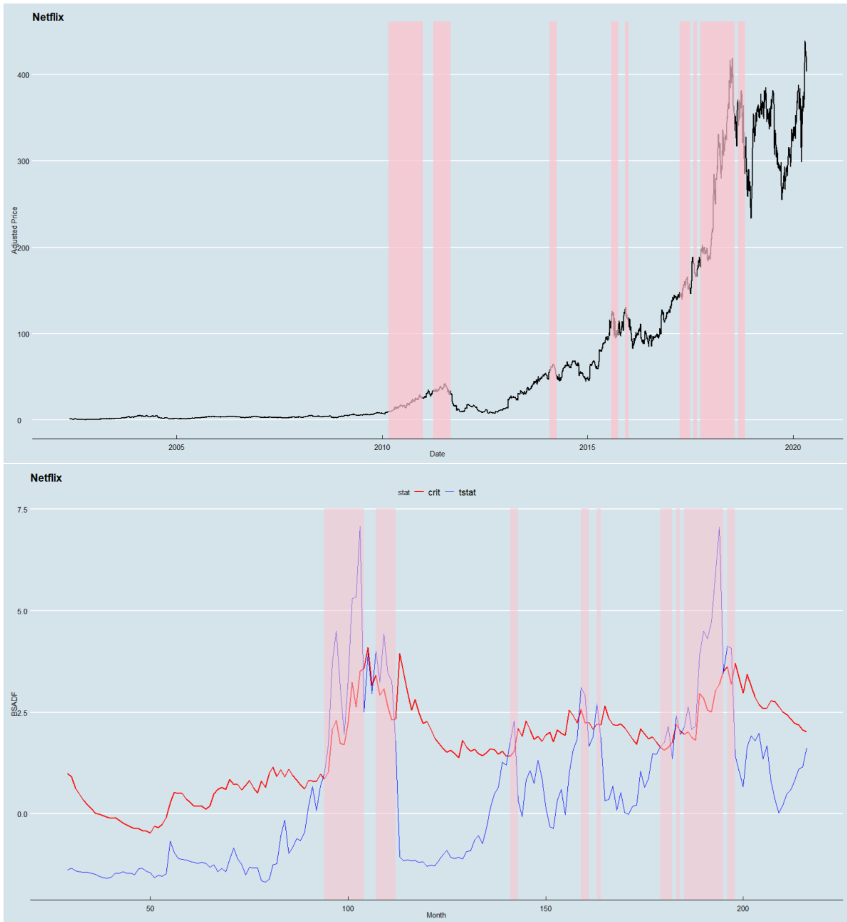

Fig. 17 In the image above, we noticed the adjusted daily price (black line) of Netflix Inc. since May 23, 2002, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Walt Disney Company according to the GSADF test, with 95% confidence intervals in Fig. 18. For this stock, we identified seven possible bubbles:

1. From August 30, 2002 to September 30, 2002.
2. From January 31, 2007 to February 28, 2007.
3. From March 28, 2013 to August 30, 2013.
4. From September 30, 2013 to August 31, 2015.
5. From October 30, 2015 to December 31, 2015.

6. From March 31, 2017 to May 31, 2017.
7. From July 31, 2019 to August 30, 2019.

Fig. 18 In the image above, we noticed the adjusted daily price (black line) of Walt Disney Company since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Adobe Inc. according to the GSADF test, with 95% confidence intervals in Fig. 19. For this stock, we identified six possible bubbles:

1. From November 30, 2004 to January 31, 2005.
2. From October 31, 2013 to April 30, 2014.
3. From October 30, 2015 to January 29, 2016.
4. From May 31, 2016 to June 30, 2016.

5. From August 31, 2016 to November 30, 2016.
6. Beginning in December 30, 2016 but has not ended yet.

Fig. 19 In the image above, we noticed the adjusted daily price (black line) of Adobe Inc. since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for PayPal according to the GSADF test, with 95% confidence intervals in Fig. 20. For this stock, we identified three possible bubbles:

1. From May 31, 2017 to February 28, 2018.
2. From April 30, 2019 to May 31, 2019.
3. From June 28, 2019 to July 31, 2019.

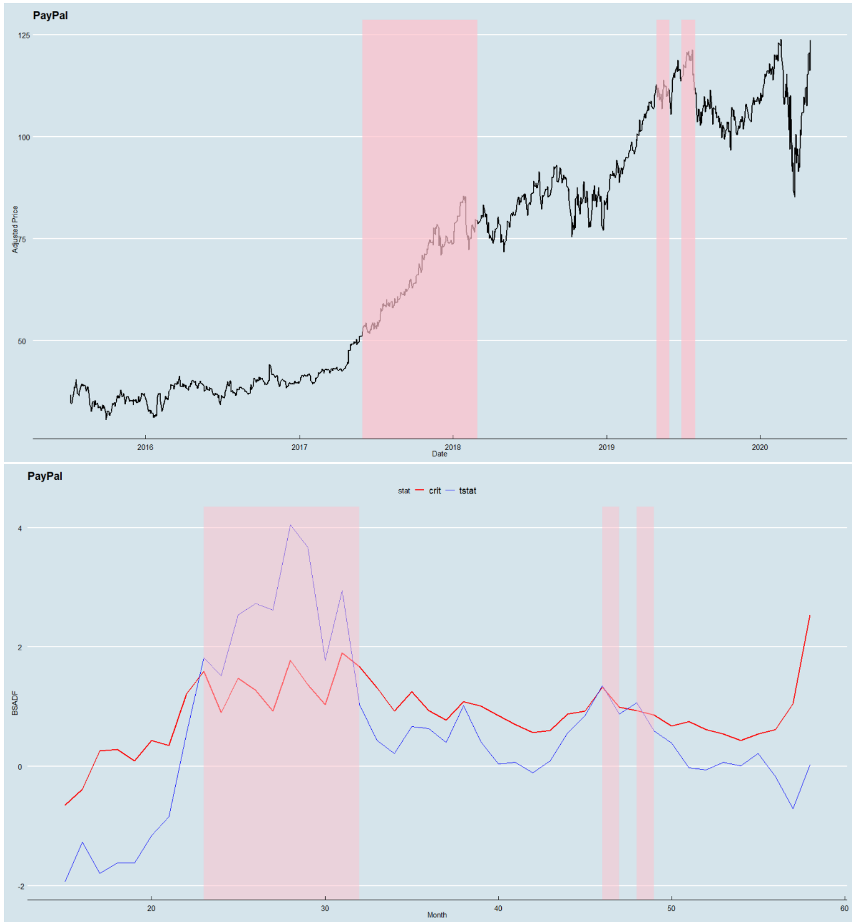

Fig. 20 In the image above, we noticed the adjusted daily price (black line) of PayPal Holdings Inc. since July 6, 2015, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for NexEra Energy Inc. according to the GSADF test, with 95% confidence intervals in Fig. 21. For this stock, we identified twelve possible bubbles:

1. From August 31, 2004 to September 30, 2004.
2. From November 30, 2004 to October 31, 2005.
3. From October 31, 2006 to June 29, 2007.
4. From September 28, 2007 to January 31, 2008.
5. From July 31, 2012 to August 31, 2012.
6. From March 28, 2013 to May 31, 2013.
7. From January 31, 2014 to July 31, 2014.

8. From November 28, 2014 to April 30, 2015.
9. From May 29, 2015 to June 30, 2015.
10. From January 29, 2016 to November 30, 2016.
11. From February 28, 2017 to March 31, 2017.
12. Beginning in April 28, 2017 but has not ended yet.

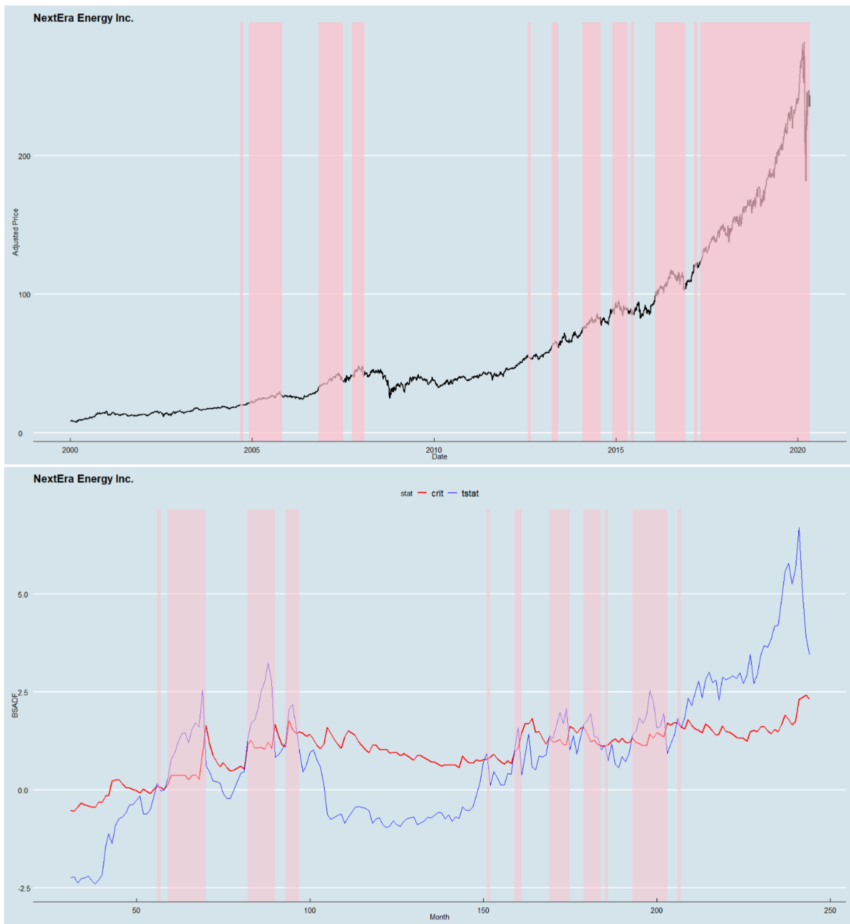

Fig. 21 In the image above, we noticed the adjusted daily price (black line) of NextEra Energy Inc. since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for NIKE according to the GSADF test, with 95% confidence intervals in Fig. 22. For this stock, we identified eleven possible bubbles:

1. From December 31, 2003 to April 30, 2004.

2. From September 30, 2004 to March 31, 2005.
3. From November 30, 2006 to June 30, 2008.
4. From February 28, 2011 to March 31, 2011.
5. From January 31, 2012 to June 29, 2012.
6. From March 28, 2013 to March 31, 2014.
7. From May 30, 2014 to September 30, 2016.
8. From February 28, 2017 to March 31, 2017.
9. From April 30, 2018 to October 31, 2018.
10. From February 28, 2019 to May 31, 2019.
11. From December 31, 2019 to January 31, 2020.

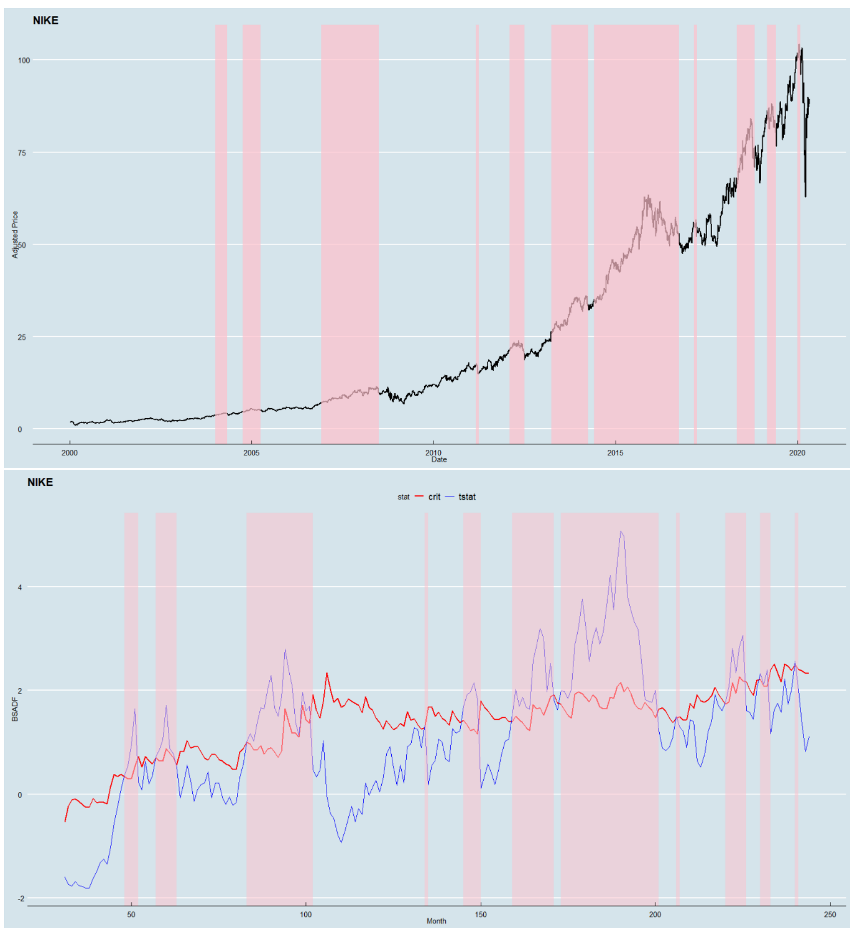

Fig. 22 In the image above, we noticed the adjusted daily price (black line) of NIKE Inc. Class B since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Union Pacific Corporation according to the GSADF test, with 95% confidence intervals in Fig. 23. For this stock, we identified thirteen possible bubbles:

1. From January 31, 2006 to February 28, 2006.
2. From March 31, 2006 to July 31, 2006.
3. From January 31, 2007 to February 28, 2007.
4. From March 30, 2007 to August 31, 2007.
5. From October 31, 2007 to June 30, 2008.
6. From July 31, 2008 to September 30, 2008.
7. From March 31, 2011 to August 31, 2011.
8. From June 29, 2012 to June 30, 2015.
9. From December 29, 2017 to October 31, 2018.
10. From November 30, 2018 to December 31, 2018.
11. From April 30, 2019 to May 31, 2019.
12. From July 31, 2019 to August 30, 2019.
13. From December 31, 2019 to January 31, 2020.

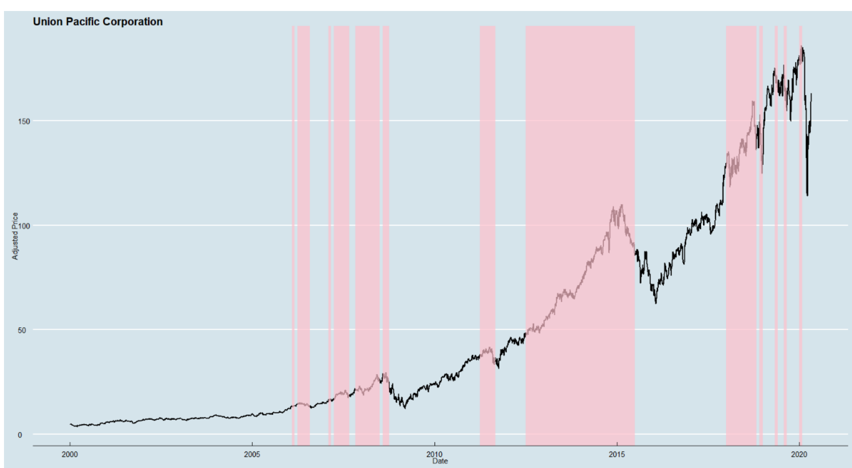

Fig. 23 In the image above, we noticed the adjusted daily price (black line) of Union Pacific Corporation since January 1, 2000, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We graph the estimated result for Broadcom Inc. according to the GSADF test, with 95% confidence intervals in Fig. 24. For this stock, we identified four possible bubbles:

1. From June 30, 2011 to July 29, 2011.
2. From February 28, 2014 to June 30, 2015.
3. From August 31, 2016 to March 29, 2018.
4. From March 29, 2019 to May 31, 2019.

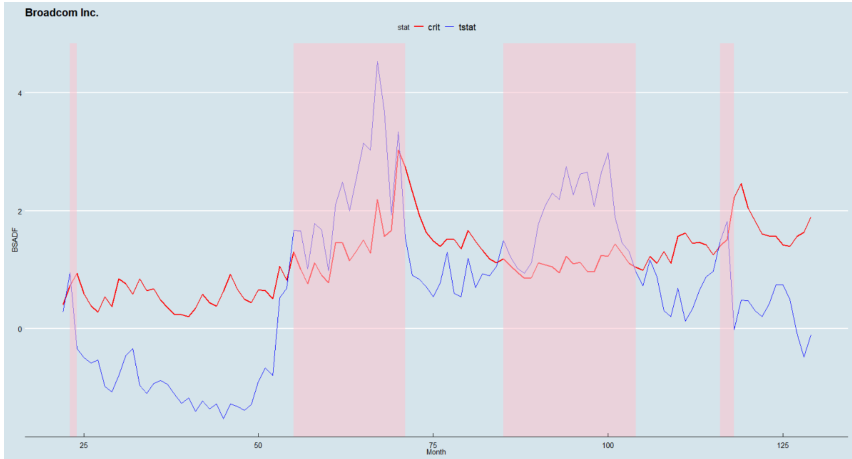

Fig. 24 In the image above, we noticed the adjusted daily price (black line) of Broadcom Inc. since August 6, 2009, with its date-stamp extracted from the GSADF. In the image below, we observed the critical values as wild bootstrap simulations (red line) and the forward Augmented Dickey-Fuller (ADF) sequence (blue line).

We summarized the above results in Table 7., where denoted the following information:

1. Company: the name of the company we studied.
2. Initial Date: the start date of the information extracted from Yahoo Finance.
3. SADF: in case we rejected the null hypothesis of the SADF test, we assigned a “Y” label to the row.
4. Bubble ID: the number of bubble-type behavior we found applying the GSADF test for each stock.
5. Start and End: both columns show the birth and collapse of each exuberance found.
6. Duration: the number of months between the Start and End columns.
7. GSADF / significance level: the rejection of the null hypothesis with a significance level of 1% or 5%.

It is noteworthy that we selected monthly data because we wanted to study the S&P 500 index sample according to the minimum time interval required to consider a “short bubble.” Additionally, Gerlach, Demos, and Sornette (2019) analyzed the price trajectory of “bubbles” and determined that this short-term behavior’s minimum duration should exceed thirty days.

Table 7. GSADF test summary for fifty stocks of the S&P 500 index.

Company	Initial Date	SADF	Bubble ID	Start	End	Duration	GSADF / significance level
Microsoft Corporation	03/01/2000	Y	1	31/07/2014	30/01/2015	6	1%
			2	31/01/2017	NA	40	
Apple Inc.	N						
Amazon.com Inc.	03/01/2000	Y	1	30/06/2003	27/02/2004	8	1%
			2	31/12/2010	31/01/2011	1	
			3	29/04/2011	30/11/2011	7	
			4	28/09/2012	31/10/2012	1	
			5	28/02/2013	30/04/2013	2	
			6	31/05/2013	30/08/2013	3	
			7	30/09/2013	31/03/2014	6	
			8	30/06/2015	31/08/2015	2	
			9	30/09/2015	29/01/2016	4	
			10	31/05/2016	31/10/2018	29	
Facebook Inc. Class A	N						
Alphabet Inc. Class A	N						
Alphabet Inc. Class C	N						
Johnson & Johnson	N						
Berkshire Hathaway Inc. Class B	N						
Visa Inc. Class A	03/01/2000	Y	1	29/02/2012	31/05/2012	3	5%
			2	31/07/2012	30/04/2014	21	
			3	30/05/2014	31/08/2015	15	
			4	30/10/2015	29/01/2016	3	
			5	31/03/2016	30/06/2016	3	
			6	31/08/2016	30/11/2016	3	
			7	28/02/2017	31/12/2018	22	
			8	31/01/2019	31/03/2020	14	
Procter & Gamble Company	N						
UnitedHealth Group Incorporated	03/01/2000	Y	1	27/02/2004	30/04/2004	2	5%
			2	30/11/2004	31/03/2006	16	
			3	30/06/2011	29/07/2011	1	
			4	30/03/2012	30/04/2012	1	
			5	31/03/2014	30/04/2014	1	
			6	31/10/2014	28/02/2019	52	
JPMorgan Chase & Co.	N						
Intel Corporation	N						
Home Depot Inc.	03/01/2000	Y	1	29/02/2012	31/10/2016	56	5%
			2	30/11/2016	31/10/2018	23	
			3	30/04/2019	31/05/2019	1	

			4	28/06/2019	31/03/2020	9	
Mastercard Incorporated Class A	25/05/2006	Y	1	31/08/2011	30/09/2011	1	5%
			2	30/11/2011	31/01/2012	2	
			3	30/03/2012	31/05/2012	2	
			4	30/11/2012	31/03/2015	28	
			5	30/04/2015	29/01/2016	9	
			6	28/04/2017	31/12/2018	20	
			7	28/02/2019	31/03/2020	13	
Verizon Communications Inc.	N						
Pfizer Inc.	N						
AT&T Inc.	N						
Merck & Co. Inc.	N*						
NVIDIA Corporation	03/01/2000	Y	1	28/02/2006	31/05/2006	3	5%
			2	31/08/2007	30/11/2007	3	
			3	30/11/2015	29/01/2016	2	
			4	31/05/2016	31/10/2018	29	
Netflix Inc.	23/05/2002	Y	1	26/02/2010	31/12/2010	10	5%
			2	31/03/2011	31/08/2011	5	
			3	31/01/2014	31/03/2014	2	
			4	31/07/2015	30/09/2015	2	
			5	30/11/2015	31/12/2015	1	
			6	31/03/2017	30/06/2017	3	
			7	31/07/2017	31/08/2017	1	
			8	29/09/2017	31/07/2018	10	
			9	31/08/2018	31/10/2018	2	
Walt Disney Company	03/01/2000	Y	1	30/08/2002	30/09/2002	1	5%
			2	31/01/2007	28/02/2007	1	
			3	28/03/2013	30/08/2013	5	
			4	30/09/2013	31/08/2015	23	
			5	30/10/2015	31/12/2015	2	
			6	31/03/2017	31/05/2017	2	
			7	31/07/2019	30/08/2019	1	
Cisco Systems Inc.	N						
PepsiCo Inc.	N						
Exxon Mobil Corporation	N						
Bank of America Corp	N						
Waltmart Inc.	N						
Adobe Inc.	03/01/2000	Y	1	30/11/2004	31/01/2005	2	1%
			2	31/10/2013	30/04/2014	6	
			3	30/10/2015	29/01/2016	3	
			4	31/05/2016	30/06/2016	1	

			5	31/08/2016	30/11/2016	3	
			6	30/12/2016	NA	41	
Chevron Corporation	N						
PayPal Holdings Inc.	06/07/2015	Y	1	31/05/2017	28/02/2018	9	5%
			2	30/04/2019	31/05/2019	1	
			3	28/06/2019	31/07/2019	1	
Coca-Cola Company	N						
Comcast Corporation Class A	N						
Abbott Laboratories	N						
AbbVie Inc.	N						
Bristol-Myers Squibb Company	N						
salesforce.com Inc.	N						
Amgen Inc.	N						
Thermo Fisher Scientific Inc.	N						
Eli Lilly and Company	N*						
Costo Wholesale Corporation	N*						
McDonald's Corporation	N*						
Medtronic Plc	N						
Oracle Corporation	N						
Accenture Plc Class A	N						
NextEra Energy Inc.	03/01/2000	Y	1	31/08/2004	30/09/2004	1	1%
			2	30/11/2004	31/10/2005	11	
			3	31/10/2006	29/06/2007	8	
			4	28/09/2007	31/01/2008	4	
			5	31/07/2012	31/08/2012	1	
			6	28/03/2013	31/05/2013	2	
			7	31/01/2014	31/07/2014	6	
			8	28/11/2014	30/04/2015	5	
			9	29/05/2015	30/06/2015	1	
			10	29/01/2016	30/11/2016	10	
			11	28/02/2017	31/03/2017	1	
			12	28/04/2017	NA	37	
NIKE Inc. Class B	03/01/2000	Y	1	31/12/2003	30/04/2004	4	5%
			2	30/09/2004	31/03/2005	6	
			3	30/11/2006	30/06/2008	19	
			4	28/02/2011	31/03/2011	1	
			5	31/01/2012	29/06/2012	5	
			6	28/03/2013	31/03/2014	12	
			7	30/05/2014	30/09/2016	28	
			8	28/02/2017	31/03/2017	1	

			9	30/04/2018	31/10/2018	6	
			10	28/02/2019	31/05/2019	3	
			11	31/12/2019	31/01/2020	1	
Union Pacific Corporation	03/01/2000	Y	1	31/01/2006	28/02/2006	1	5%
			2	31/03/2006	31/07/2006	4	
			3	31/01/2007	28/02/2007	1	
			4	30/03/2007	31/08/2007	5	
			5	31/10/2007	30/06/2008	8	
			6	31/07/2008	30/09/2008	2	
			7	31/03/2011	31/08/2011	5	
			8	29/06/2012	30/06/2015	36	
			9	29/12/2017	31/10/2018	10	
			10	30/11/2018	31/12/2018	1	
			11	30/04/2019	31/05/2019	1	
			12	31/07/2019	30/08/2019	1	
			13	31/12/2019	31/01/2020	1	
Broadcom Inc.	06/08/2009	Y	1	30/06/2011	29/07/2011	1	5%
			2	28/02/2014	30/06/2015	16	
			3	31/08/2016	29/03/2018	19	
			4	29/03/2019	31/05/2019	2	
Phillip Morris International Inc.	N						
International Business Machines Corporation	N						

On the one hand, and just as a reminder, for the stocks labeled with an “N” we could not reject the null hypothesis of the SADF test. In contrast, the stocks labeled with a “N*” rejected the null hypothesis of the SADF test but not the generalized sup Augmented Dickey-Fuller test (GSADF). On the other hand, the maximum duration of exuberance found in the S&P 500 index sample was 56 months and belonged to Home Depot (HD), from February 29, 2012, to October 31, 2016, while the average duration was of 8.4 months. Plus, it is essential to mention that the marker “NA” in the column named “End” indicates the current explosive behavior has still not collapsed.

Furthermore, we realized it was necessary to visualize the interaction of the 106 “bubbles” found in a graph, so in Fig. 25, we established each exuberance for the fifteen stocks that rejected the null hypothesis of the GSADF test. In other words, we showed the links between the "bubbles" across the S&P 500 index. Also, in Table 8, we showed the assets' sector with at least one financial bubble detected.

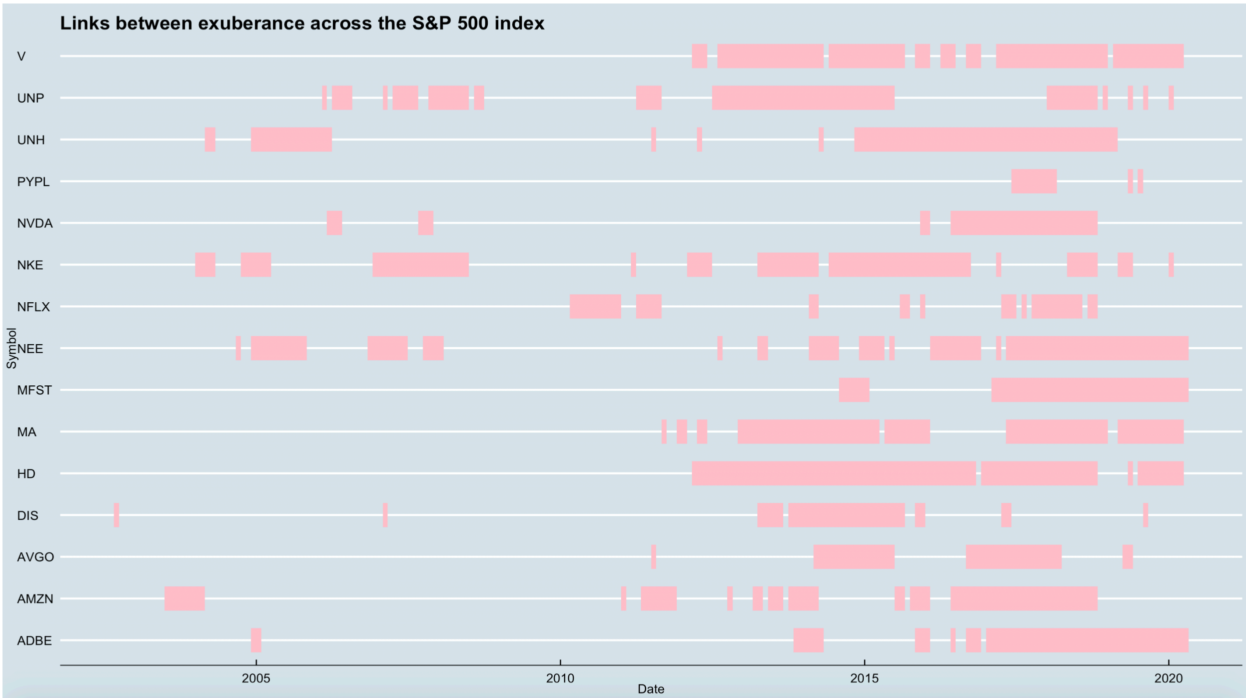

Fig. 25 Exuberance periods defined according the GSADF in fifteen stocks of the S&P 500 index.

Table 8. Sector breakdown.

Sector breakdown	Quantity	Percentage
Communication Services	2	13.33%
Consumer Discretionary	3	20.00%
Health Care	1	6.67%
Industrials	1	6.67%
Information Technology	6	40.00%
Utilities	2	13.33%
Total	15	100.00%

We identified twenty-two "bubbles" that ended before December 31, 2010, and eighty-four that occurred after or still do not finish until April 29, 2020. Hence, we were able to detect that the identified financial bubbles belong in a forty percent to the Information Technology sector, and in second place, the Consumer Discretionary takes place. Consequently, the exuberance evolution in the XXI century has increased in some stocks of the S&P index where NextEra Energy Inc, NIKE Inc. Class B, and Union Pacific Corporation have more than ten “bubbles.”

3.5 CONCLUSION

In conclusion, we have presented an analysis of fifty stocks of the S&P 500 index from January 3, 2000, to April 29, 2020, where for fifteen of them, we were able to find at least more than one bubble-type behavior according to the generalized sup Augmented Dickey-Fuller (GSADF) test. Given the results and the ascent appearance of exuberance in one of the most important stock market indices in the United States and the world, we notice an almost uninterrupted regime where the markets keep growing since the end of the 2008 crisis. Therefore, in more than a decade, financial assets have been increasing their demand at higher values, sky-rocketing price performance, and possibly going beyond their fundamental value in many cases. We only examined fifty stocks of the entire index; despite the selected components that constituted around 50 percent of the index, we found 30 percent of the sample with at least one “bubble” in the XXI century. Consequently, the exuberant trend could be in the entire market.

Intentionally left blank

4. CHAPTER III: NORMAL INVERSE GAUSSIAN EPISODES

4.1 INTRODUCTION

In retrospect, macroeconomic theorists were the first to define financial bubbles' possible existence or describe them as "growing bubbles." Camerer (1989) showed an explosive solution in dynamic models even if agents have rational expectations. "The short story of formal theory about growing bubbles is an intellectual struggle between attempts to rationalize the possibility of bubbles, because they may occur, and attempts to rule our bubbles because they are arbitrary" (p. 7).

The standard definition of financial bubbles is a rapid price increase that deviates from its fundamental value. They are apparently observable conditions, but it is clear that their detectability is not a straight forward job. The previous chapter showed a statistic methodology to identify explosive behaviors, where financial assets are appropriately included. In other words, the generalized sup Augmented Dickey-Fuller (GSADF) detects multiple bubbles (Phillips et al., 2015); however, it does not give us more information about the identified exuberance. Therefore, in this chapter, we analyzed the implications of financial bubbles regarding their fit to the Normal Inverse Gaussian (NIG). This section is based on our previous research (Núñez, Contreras-Valdez, & Franco-Ruiz, 2019); nevertheless, we extended the research to the asset price bubbles identified in the previous chapter.

4.2 THEORETICAL FRAMEWORK

Blanchard and Watson (1982) refer to leptokurtosis as a general characteristic of financial returns. Moreover, it seems widely accepted in the literature that leptokurtosis is a characteristic of speculative bubbles (Meese, 1986; Camerer, 1989). Nevertheless, Evans (1986) mentioned the statistical property derived from kurtosis, autocorrelation, and a non-zero median in the distribution of price changes due to market fundamentals rather than a bubble-type behavior. We should not only focus on the kurtosis because it is a limited deviation measure from

Gaussian statistics; well-known returns' distribution is leptokurtic and fat-tailed[5] (Lux and Sornette, 2002).

Koedijk, Schafgans, and De Vries (1990) specified that the appropriate distribution would depend on the degree of tail-fatness. Hence, the normal distribution will be dismissed immediately due to "the excessive amount of outliers in the data." (p. 94). Additionally, they applied extreme value theory as a way to explain the next three problems:

1. The class of distribution function.
2. Parameters consistency over subsamples.
3. Distribution effects of aggregation over time.

Jansen and De Vries (1991) investigated the tail behavior instead of the entire distribution of stock returns to find explanations about the probability mass in the tails. They employed an extreme value theory approach. Consequently, Loretan and Phillips (1994) researched methods for testing the assumption that time series' unconditional variance is constant over time because they observed that many financial datasets have heavy tails. They raised the question of how volatility should be studied and the appropriate methods to model volatility.

Meanwhile, Pagan (1996) elaborated holistic research on financial econometrics. He concentrated on three types of series; stock prices, interest rates, and exchange rates, given that "financial data appears in many forms…", where he made emphasis that each time series has "its own idiosyncrasies." (p. 16).

Eberlein and Keller (1995) mentioned that the "normal distribution is a poor model for stock returns," unlike the hyperbolic distributions that can fit with high accuracy (p. 284). Eberlein and Prause (2002) mentioned the results concerning the

[5] It refers to the tails' probability weight, so we mean by fat-tailed distributions to distributions that have fatter tails than the normal distribution.

GH, hyperbolic, and the normal inverse Gaussian (NIG) are much closer to the empirical distribution, so the generalized hyperbolic distributions “seem to be tailor-made to describe the statistical behaviour of asset returns.” (p. 245). Therefore, we can use the GH distribution and its subclasses to calculate the Value-at-Risk (VaR) and obtained a closer empirically potential loss given a level of probability. Rydberg (1999) showed that the NIG diffusion process has a good fit for major US stocks log returns.

Cont (2001) stated:

> “… the result of more than half a century of empirical studies on financial time series indicates that this is the case if one examines their properties from a statistical point of view: the seemingly random variations of asset prices do share some quite non-trivial statistical properties. Such properties, common across a wide range of instruments, markets and time periods are called stylized empirical facts.” (p. 224)

Hence, we should understand stylized facts like common factors or properties observed in diverse financial markets, where Cont (2001) describe the following properties:

1. Absence of autocorrelations
2. Heavy tails
3. Gain/loss asymmetry
4. Aggregational Gaussianity
5. Intermittency
6. Volatility clustering
7. Conditional heavy tails
8. Slow decay of autocorrelation in absolute returns
9. Leverage effect

10. Volume/volatility correlation
11. Asymmetry in time series

Also, he mentioned some issues about statistical estimation, which are "implicit in almost any statistical analysis of asset returns." (Cont, 2001, p.224). Those issues are stationarity, ergodicity, and finite sample properties of estimators.

Among many models that have been scrutinized, the class of hyperbolic distributions is an excellent candidate besides the normal distribution. Barndorff-Nielsen (1977) introduced this class of distribution, where we can determine the generalized hyperbolic (GH) distribution by five parameters to fit empirical distributions to financial assets. The following paragraphs make a quick overview of several applications of NIG in financial and actuarial sciences.

Fosberg and Bollerslev (2002) find volatilities constructed with summations of squared returns conditioned on lagged squared daily returns are "inverse Gaussian" distributed, with a normal distribution for volatility – standardized daily returns. They prove that the out-of-sample prediction for three years, using a NIG adjusted daily GARCH model, fit Euro/Dollar exchange rates well. The GARCH NIG model is recently introduced to literature and proves to provide an accurate representation of rate dynamics.

Boyarchenko and Levendorskii (2002) derive formulas for European barrier options and touch-and-out options considering that under an equivalent martingale measure, stock returns follow a Lévy process, considering the presence of Brownian Motions, NIG processes, hyperbolic processes, and other finite mixtures of the processes mentioned. This development comes in aid in the definition of pricing formulas.

Karlis and Lillestöl (2004) propose the estimation of NIG parameters via Markov chain Monte Carlo scheme based on the Gibbs algorithm for estimating a sample from a multivariate probability distribution, under a Bayesian estimation scheme. Albrecher and Predota (2004) compare NIG average option prices with the corresponding B-S prices, considering Asian European style average options (arithmetic and geometric). They assume an exponential Levy process for the price of assets and NIG distributed logarithmic returns. The study shows accurate approximations between the pricing methods. The use of the Esscher equivalent measure is an important assumption made in the study, and the authors justify its use according to utility and equilibrium theory (Gerber and Shiu, 1994). The presence of the Esscher equivalent measure in the NIG model offers a simple structure that is exploited to obtain approximations for geometric and arithmetic average option pricing. One observation that proves the simplicity of Esscher's price calculation under NIG model is that the Esscher transform of a random variable, which is NIG distributed, is again NIG distributed. They conclude their study with a comparison of prices, showing that the NIG approximation outperforms the Turnbull-Wakeman approximation and the Lévy approximation.

Rasmus, Asmussen, and Wiktorsson (2004) study Russian and barrier option pricing problems considering the assumption of exponential NIG and Lévy processes through simulation. They address the issue simulating discrete grids and the neglection of minimum and maximum between grid points. They propose the simulation of large jumps and using a Brownian motion approximation for the rest of jumps, using formulas for minima and maxima.

Chang et al. (2005) test the symmetry for a NIG distribution proposing a Likelihood Ratio Test. They use the EM (Expectation-Maximization) type algorithm (Karlis, 2002) to find the distribution estimates. One impressive result is that the testing power for the NIG increases as its asymmetry increases. This development's primary conclusion is that the proposed likelihood ratio test proved to detect asymmetric behavior.

Venter, Jongh, and Griebenow (2005) develop an extension to the Lindhodt's (2002) model, introducing the NIG approach to fitting a GARCH model for financial return series open, close, high, and low prices are used. Since the normal innovation distribution for the GARCH model produces light tails for the return process, the use of NIG means to amend the need for a more realistic approach.

Kassberger and Kiesel (2006) suggest a model using the multivariate version of the NIG distribution to capture some features of hedge funds returns. They intend to develop a risk management framework using parameters of Generalized Hyperbolic distributions, aiming to describe the properties of univariate and multivariate returns, given the characteristics of hedge funds performance (non-normality of returns). Given that EVT methodologies and VaR estimations tend to use the tail data and discard the rest of the observations, the calibration process becomes difficult due to the few observations available. To address this problem, the authors advocate using NIG distribution, a model that does not suffer from the same shortcomings as EVT and VaR modeling. It is demonstrated that this asset class shows pronounced skewness and excess kurtosis, making the GH – NIG distribution suitable for these returns.

Benth, Groth, and Kettler (2006) propose using a Monte Carlo method based on the sampling of three uniform independent variables to simulate variates from the NIG distribution to create a quasi-Monte Carlo algorithm. Their study considers valuation for vanilla call options and Asian options and analyzes the underlying dynamics of assets based on the observer option prices. They use the algorithm to determine the NIG scale parameter for log-returns distribution of Asian options and evaluate the Value at Risk estimation for a non-linear portfolio with NIG random variables modeled returns.

Albrecher, Ladoucette, and Schoutens (2007) bring up the weakness of one factor Gaussian models in fitting prices for several synthetic CDO tranches. Given the one-factor model, including different distributions, the authors unify the approaches using NIG, Variance-Gamma, and Brownian Variance Gamma. One-factor models describe one factor Lévy model, applied to work a sizeable homogeneous portfolio. This approximation is used to calculate the tranches' premium and to determine a loss distribution. The flexibility obtained from VG, NIG, and BVG allows more effective determination of dependence structures and tail dependence. Kalemanova, Schmid, and Werner (2007) develop a similar study, in which they prove the inefficiency of t-Student distribution in CDO's pricing, given that this distribution is not stable under convolution. Aside from this, the computational force required to develop large applications, especially asset allocation determination, is not easily achieved. Using NIG distribution and parameters helps improve the computational time required and creates better and more flexible dependence structures in LHP (large homogeneous portfolios) models.

Kilic (2007) proposes an extension of a fractionally integrated GARCH model with the incorporation of a NIG distribution, in order to be able to determine time variation, fat tails, and some symmetry characteristics for financial returns. Comparing GARCH and FIGARCH models for log exchange returns using normal, t-student, and NIG error distributions show that the use of NIG outperforms the other distributions, both in-sample fit and predictive ability for 1-day and 5-days forecasts.

Wilhelmsson (2009) proposes the utilization of NIG distribution for Value at Risk calculation and calibration of varying variance, skewness, and kurtosis in the model. NIG being closed under convolutions and having a closed-form density makes it a suitable choice for fitting data to the distribution. In the paper, the author compares NIG – ACD (Autoregressive conditional density) and NIG-S&ARCH (Stochastic and Autoregressive conditional heteroskedastic) volatility models (Jensen and Lunde, 2001), having the first one providing enough independent VaR exceptions in six levels evaluated. According to Basel rules, NIG based models accurately determine capital requirement, meaning that there is no need for further calibration nor additional capital requirements. NIG – ACD outperforms NIG-S & ARCH models in-sample fitting and out sample density forecasts and VaR estimates.

The author identifies the importance of the conditional variance, conditional skewness, and time-varying kurtosis estimation parameters gained with this model. Given that the standard's industry for risk modeling and VaR calculation uses GARCH-n models, these results indicate the need to actualize the requirements and methodologies defined in Basel.

Frestad, Benth, and Koekebaker (2010) analyze Nordic electricity swaps and returns, identifying risk premia, negative return for short positions, and non-normal daily returns. They test the four-parameter NIG distribution to determine the model's ability to observe stylized facts and a better fitting compared to normal distribution. This study proves again that NIG applications come useful for pricing derivatives and measuring VaR estimates. The authors compare stable distributions with NIG, and they find out that NIG law outperforms the stable law for most cases.

Jeannin and Pistorious (2010) propose computing barrier options prices and Greeks when these options show to be driven by a hyper-exponential Lévy process. Considering several other approaches applied for barrier option prices and some sensitivities, they adopt a CGMY Lévy density and the Wiener Hopf factorization to derive analytical formulas for Laplace transform when knocked in or out option prices. With such results, the authors also derive the sensitivities delta, gamma, and theta. Prices and greeks are obtained by inverting Laplace transformations employing Abate and White's algorithm. The numerical illustration included in the study implements this algorithm for VG and NIG Lévy models. Compared to Monte Carlo, the relative error rate was around 0.5%-2.5%.

Hainaut and Macgilchrist (2010) propose the derivation of an interest rate model driven by a NIG process instead of the Brownian Motion. NIG can capture excess kurtosis and skewness, typical of interest rates distribution nature. Better fitting for bond returns and better capturing asymmetry and leptokurtic of short-term rates are additional motivations for choosing the NIG. The authors compare this development with the Hull and White model, except for the NIG process replacing the BM. They calibrate for prices using a pentanomial tree to reproduce the four moments of the NIG mean-reverting process. The numerical tests for NIG application outperform the White-Hull model with a significant margin. It also provided a new characterization of a parameter, distinguishing curve shapes, and fitting steep volatility curves.

Godin, Mayoral, and Morales (2012) address the contingent claim pricing problem employing distorted operators. This methodology was first applied by Wang (2000) in insurance risk pricing, using normal distortion risk “quantile-based” measures. The distortion operation effectively derives the Black Scholes formula, considering normal distribution for asset prices (Hamada and Sherris, 2003). Given the nature of financial assets returns, and the presence of asymmetry and fat tails. Several studies prove that non-gaussian distributions, specifically those from Generalized Hyperbolic family, adjust better to these qualities. The authors propose a distortion operator based on NIG distribution and acquire the option formula within a mean-correcting class of equivalent martingale measures. They also develop a simulation exercise to illustrate how NIG distortion operator shows improved robustness compared with other asset price models that work with normal based and student operators.

Hofer and Mayer (2013) address the problem-pricing lookback options for exponential Lévy models, using Laplace transforms and the fist passage time distribution for Hyper Exponential Jump Diffusion processes. In the numerical results of this investigation, the authors derive prices and sensitivities for the lookback options using the Gaver-Stehfest algorithm for the inversion of Laplace transform and compare the results with Monte-Carlo approximations for the HEJD process. They fit a NIG process to the HEJD process and compare the prices with a Monte Carlo simulation of the NIG process.

Von Hammerstein et al. (2014) take Levy models for prices driven by NIG and VG processes to determine low-cost strategies for given payoffs in a Lévy market, where prices are based on the Esscher martingale measure. Authors compare NIG and VG processes to the Black-Scholes model, and derive cost-efficient strategies for path independent payoffs; provide a cost-efficient version of put-call parity; and determine hedging strategies for cost-efficient payoffs, providing delta hedging formulas for the cost-efficient strategies for European call and put option. They demonstrate that all formulas derived are tractable, and hedging strategies can be determined accurately. They provide numerical examples that prove the improvement acquired by switching to the cost-efficient strategy developed in this paper. Also, prove that efficiency losses modeled through different Lévy processes produce similar magnitudes, defining that results are model-independent.

Yamazaki (2014) presents an approximation of exotic option pricing for underlying asset prices driven by time changed Lévy processes. Time-change processes are considered useful for capturing random time changes effect on stochastic volatility, a flexible jump generating framework, and leverage effect. By a Gram Charlier expansion, the author derives a pricing formula that produces accurate approximations for average option prices, with efficient computational effort. By this, they avoid the use of Monte Carlo simulation, as it is time-consuming and tends to produce inaccurate estimations. The author adopts the Heston Model and a VG-CIR and NIG-CIR to describe the asset price dynamics for the numerical examples. NIG-CIR takes the normal inverse gaussian process as background Lévy process, and the CIR process for capturing the activity rate process of the time change, with 6 to 7 iterations, the models can accurately approximate the prices for average continuously monitored and discretely monitored call options.

Chuang and Brockett (2014) take the Lee-Carter mortality model improvements made by MBMM (Mitchell et al. 2013), where mortality rates growth rates are fitted over time and age instead of the growth rate itself. The authors unify these two concepts and models for the determination of longevity/mortality linked derivatives. Given that mortality rates are not traded themselves, the derivative must be priced on an incomplete market, so a stochastic model for mortality rate must be adopted to determine equivalent martingale measures. The authors propose to model the mortality rate growth with Mitchel's MBMM model and link this model with martingale pricing; they introduce a Lévy process for the dynamic mortality modeling. They adapt the MBMM modeling for the martingale pricing measure using a NIG distribution specification and apply an Esscher transform to obtain the measure, capturing jumps and kurtosis and skewness in mortality growth rates. Results are compared to the LifeMetrics index and show that in sample estimations can reproduce prices under the mentioned model. Another remark is that the model derives higher forward rates, lower premiums, and reduced hedging costs for longevity risk.

Kirkby (2017) presents a method for discretely monitored barrier options and occupation time derivatives pricing using exponential Lévy models. It is essential to highlight that there are no closed formulas for analytical valuation for this class of path-dependent exotic contracts. The authors make use of a backward induction method, which is based on the projection frame approach. Compared to the COS method, non-orthogonal basis expansions are considered altogether with a biorthogonal basis for determining the orthogonally projected density. The authors devise an automated parameter selection method to enable pricing without the user's intervention and facilitate direct calibration, so the grid size and tolerance meeting are automatically determined. CPU times for NIG and other Lévy processes outperformed the computational efficiency compared to the COS method. They demonstrate the convergence of their model through several numerical approximations and using a Toeplitz representation of intermediate value coefficients, and they derive accurate prices for the derivates.

Cao et al. (2020) determine the modeling for VIX derivatives via a two-factor model with infinite-activity jumps to reduce pricing errors. VIX has two methodologies for pricing; from instantaneous volatility (usually considering finite-activity jumps as well), the alternative methodology determines the VIX dynamics and, in a second process, the pricing formulae for derivatives prices, considering stochastic volatilities and long term mean reversion. For the jump process, authors select the VG and NIG processes; VG for finite variation and NIG for infinite variation. To obtain an accurate model calibration, they combine the UKF (Unscented Kalman Filter) and QMLE (Quasi Maximum Likelihood Estimation) joint estimation approaches on the VIX and VIX derivatives data. Several two-factor models were put to the test in order to determine the most efficient configuration. Results indicate that infinite activity jump models perform better given VIX derivatives' nature and the pricing process, mostly when high-frequency data is involved. It is also noted that the two-factor model OU-NIG is better than the model developed with a Variance Gamma process factor. The NIG factor model gave better results for short maturity derivatives, Lévy jump structures suit to the high-frequency occurrences and small jump events that compose the structure of VIX derivatives.

Barndorff-Nielsen (1977) defined the NIG mathematically as follows:

$$g(x;\alpha,\beta,\mu,\delta) = a(\alpha,\beta,\mu,\delta)q\left(\frac{x-\mu}{\delta}\right)^{-1} K_1\left\{\delta\alpha q\left(\frac{x-\mu}{\delta}\right)\right\}\exp\left(\beta x\right) \tag{1}$$

where,

$$a(\alpha,\beta,\mu,\delta) = \pi^{-1}\alpha exp\left(\delta\sqrt{\alpha^2-\beta^2}-\beta\mu\right) \tag{2}$$

where,

$$q(x) = \sqrt{1+x^2} \tag{3}$$

K_1 is the modified Bessel function of third order, $\alpha, \beta, \mu,$ and δ are parameters that satisfy $0 \leq \beta \leq \alpha$, $\mu \epsilon R$, and $0 < \delta$. Such that α represents the flatness of the density, which represents the values concentration around the μ (mean). β indicates the skewness and δ states the scale. It is noteworthy that in this chapter we use the original parameterization of Barndorff-Nielsen (1977) and in order to obtain the parameters that fit the empirical data, we have to apply a maximum likelihood estimation process.

The primary study object is the assets' returns. Nevertheless, as far as we know, no one has studied the fit of a theoretical distribution in periods of bubble-type behavior until our proposal that developed it through the normal inverse Gaussian (NIG) distribution on the returns of bitcoin (BTC). The NIG was able to fit every segment despite the bubble rise and collapse. Likewise, the NIG has a similar adjustment to the generalized hyperbolic (GH) distribution in the out-of-sample tests (Núñez et al., 2019).

4.3 DATA

Considering the GSADF test summary for fifty stocks of the S&P 500 explored in Chapter II Table 7, we selected the time series according to two criteria. First, it had to begin on January 3, 2000; and second, it must present bubble-type behavior in the period from January 3, 2000, to December 31, 2009 (1P), and from January 4, 2010, to April 29, 2020 (2P). Therefore, the time series that accomplished these characteristics are Amazon.com Inc., UnitedHealth Group Incorporated, NVIDIA Corporation, Walt Disney Company, Adobe Inc., NextEra Energy Inc., NIKE Inc. Class B, and Union Pacific Corporation. The next Table exhibit the descriptive statistics of each analyzed period of the assets' logarithmic daily adjusted returns mentioned above obtained from Yahoo Finance.

Table 9. Descriptive statistics summary of eight stocks of the S&P 500 index.

Asset / Period	Sample	Mean	Std. Dev	Median	Min	Max	Range	Skew	Kurtosis
AMZN 1P	2,512	0.0002	0.0400	0.0000	-0.2800	0.3000	0.5800	0.4500	8.1300
AMZN 2P	2,596	0.0011	0.0200	0.0000	-0.1400	0.1500	0.2800	0.1300	6.8300
UNH 1P	2,512	0.0006	0.0200	0.0000	-0.2100	0.3000	0.5000	0.3100	18.1500
UNH 2P	2,596	0.0009	0.0200	0.0000	-0.1900	0.1200	0.3100	-0.4700	12.7500
NVDA 1P	2,512	0.0006	0.0500	0.0000	-0.4300	0.3500	0.7900	-0.3000	10.1500
NVDA 2P	2,595	0.0011	0.0300	0.0000	-0.2100	0.2600	0.4700	0.1500	9.9500
DIS 1P	2,512	0.0001	0.0200	0.0000	-0.2000	0.1500	0.3500	0.0000	7.6000
DIS 2P	2,596	0.0005	0.0100	0.0000	-0.1400	0.1300	0.2700	-0.3100	12.6200
ADBE 1P	2,512	0.0003	0.0300	0.0000	-0.3500	0.2100	0.5700	-0.4500	8.7700
ADBE 2P	2,596	0.0008	0.0200	0.0000	-0.2100	0.1600	0.3700	-0.3700	14.2900

NEE 1P	2,512	0.0006	0.0200	0.0000	-0.1200	0.1300	0.2600	0.2200	8.3500
NEE 2P	2,596	0.0007	0.0100	0.0000	-0.1400	0.1300	0.2700	-0.3300	19.8300
NKE 1P	2,512	0.0008	0.0200	0.0000	-0.2200	0.1300	0.3500	-0.4800	10.4700
NKE 2P	2,596	0.0008	0.0200	0.0000	-0.1200	0.1400	0.2700	0.2800	11.1900
UNP 1P	2,512	0.0007	0.0200	0.0000	-0.1500	0.0900	0.2400	-0.2600	3.7200
UNP 2P	2,596	0.0007	0.0200	0.0000	-0.1400	0.1200	0.2600	-0.4200	8.3000

4.4 METHODOLOGY

For the statistical analysis, we took the results achieved in Chapter I that passed the GSADF test to select the periods in which bubble-type episodes occurred. Furthermore, we decided to separate the data into two periods (1P and 2P). The in-sample data goes from January 3, 2000, to December 31, 2009, while the out-of-the-sample test corresponds to prices from January 4, 2010, to April 29, 2020.

We investigated a unitary root in the time series at levels and proved that the original price series are non-stationary. We calculated the logarithmic returns and confirmed the series became integrated of order zero—annex II.

The procedure includes a normality test to provide evidence of the stylized facts described previously; later, two-sample goodness of fit test between empirical returns and simulated ones was performed to confirm the possibility to adjust the NIG distribution. Therefore, we can confirm, the NIG is appropriate in the return analysis and simulation when times series show explosive prices. Finally, we obtained the VaR and CVaR for the out-of-the-sample data.

In particular, from the above descriptive statistics, we obtained a non-zero skew (except DIS 1P) and an excess of kurtosis for all data, both as properties of financial series. While on the other hand, for the normality test, we chose the following criteria: Anderson-Darling, Shapiro-Francia, Lilliefors, and Cramér-von Mises.

4.5 RESULTS

In the following Table, we present the results of the normality tests.

Table 10. Normality tests.

Asset / Period	Anderson-Darling	Shapiro-Francia	Lilliefors	Cramér-von Mises	Jarque-Bera
ADBE 1P	3.70E-24	1.03E-32	6.13E-49	7.37E-10	0.000
ADBE 2P	3.70E-24	2.45E-37	4.25E-51	7.37E-10	0.000
AMZN 1P	3.70E-24	2.09E-35	2.22E-55	7.37E-10	0.000
AMZN 2P	3.70E-24	9.32E-32	2.33E-39	7.37E-10	0.000
DIS 1P	3.70E-24	7.62E-31	4.22E-33	7.37E-10	0.000
DIS 2P	3.70E-24	2.59E-37	7.73E-56	7.37E-10	0.000
NEE 1P	3.70E-24	7.41E-34	8.00E-45	7.37E-10	0.000
NEE 2P	3.70E-24	2.79E-39	1.02E-43	7.37E-10	0.000
NKE 1P	3.70E-24	8.96E-34	8.51E-64	7.37E-10	0.000
NKE 2P	3.70E-24	2.14E-36	1.58E-40	7.37E-10	0.000
NVDA 1P	3.70E-24	1.43E-33	3.38E-38	7.37E-10	0.000
NVDA 2P	3.70E-24	6.15E-34	1.41E-40	7.37E-10	0.000
UNH 1P	3.70E-24	2.32E-37	5.74E-47	7.37E-10	0.000
UNH 2P	3.70E-24	5.22E-35	9.25E-42	7.37E-10	0.000
UNP 1P	3.70E-24	3.37E-25	9.36E-34	7.37E-10	0.000
UNP 2P	3.70E-24	7.07E-31	1.01E-30	7.37E-10	0.000

We obtained the parameters for each asset for the first period of the series, and exhibit the result in the following Table.

Table 11. Adjusted parameters obtained under the maximum likelihood criteria.

Asset / Period	μ	δ	α	β
ADBE 1P	0.000317	0.022602	19.202956	0.010192
AMZN 1P	-0.000731	0.024650	14.497037	0.531170
DIS 1P	-0.000584	0.017757	35.073631	1.312626
NEE 1P	0.000958	0.011400	42.714638	-1.312578
NKE 1P	0.000143	0.014363	30.197434	1.289511
NVDA 1P	-0.000006	0.035477	16.422888	0.279613
UNH 1P	0.001008	0.016455	32.217237	-0.773279
UNP 1P	0.000469	0.016444	40.088995	0.523600

So, we can determine the distribution of the analyzed series. One of the parameters of the NIG distribution is fixed such that $\lambda = -0.5$, using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm under maximum likelihood estimation (MLE) criteria, we obtained the four parameters left by this quasi-Newton method. By comparing the parameters, we observed the highest α belongs to NEE 1P, meaning a higher concentration of probability around the mean. Therefore, we employed goodness-of-fit criteria to prove that the theoretical distribution fits the empirical data. Hence, we simulated some series for each asset and period using the previous parameters and calculated the Anderson-Darling, Kolmogorov-Smirnov, and Kruskal-Wallis statistics exposed in the next Table.

Table 12. Goodness-of-fit test (p-values).

Asset / Period	Anderson-Darling	Kolmogorov-Smirnov	Kruskal Wallis
ADBE 1P	0.92683	0.968747023	0.9754764
AMZN 1P	0.62912	0.526565215	0.33335425
DIS 1P	0.75847	0.467258304	0.58370715
NEE 1P	0.90718	0.834257475	0.87399188
NKE 1P	0.22728	0.956064393	0.36964148
NVDA 1P	0.93866	0.993496424	0.98183286
UNH 1P	0.42928	0.366926893	0.33335425
UNP 1P	0.66256	0.77858399	0.60725994

Determined by the p-values of these statistics, we cannot reject the null hypothesis; consequently, the NIG parameters could model the eight assets' observed returns.

Subsequently, as we proved that the NIG could fit periods (1P), including explosive behaviors, we can use it to estimate the VaR and CVaR. Therefore, we calculate the VaR and CVaR for the period 1P at 95%, 99%, and 99.9% confidence level. The results are shown in the next two Tables.

Table 13. VaR for NIG and GH distributions.

	NIG			GH		
Asset / Period	95%	99%	99.9%	95%	99%	99.9%
ADBE 1P	-0.051898	-0.098975	-0.182057	-0.051657	-0.098672	-0.179683
AMZN 1P	-0.060792	-0.118679	-0.218855	-0.060149	-0.119056	-0.236769
DIS 1P	-0.034540	-0.061605	-0.106628	-0.033955	-0.062036	-0.115408
NEE 1P	-0.024751	-0.047078	-0.085127	-0.024052	-0.047674	-0.097833
NKE 1P	-0.031899	-0.060621	-0.110988	-0.032085	-0.060427	-0.107936
NVDA 1P	-0.071093	-0.129003	-0.223154	-0.069216	-0.129039	-0.259065
UNH 1P	-0.034502	-0.064480	-0.115365	-0.033436	-0.064233	-0.134561
UNP 1P	-0.030864	-0.055588	-0.096611	-0.031126	-0.055564	-0.094100

Table 14. CVaR for NIG and GH distributions.

	NIG			GH		
Asset / Period	95%	99%	99.9%	95%	99%	99.9%
ADBE 1P	-0.081555	-0.133542	-0.219447	-0.081568	-0.133707	-0.220073
AMZN 1P	-0.097157	-0.161429	-0.269249	-0.097798	-0.167534	-0.293618
DIS 1P	-0.051718	-0.081101	-0.128134	-0.051788	-0.084405	-0.144409
NEE 1P	-0.038860	-0.063442	-0.103720	-0.039400	-0.068648	-0.126596
NKE 1P	-0.050140	-0.081924	-0.134355	-0.050104	-0.081022	-0.130773
NVDA 1P	-0.107777	-0.170867	-0.272490	-0.108083	-0.182923	-0.339811
UNH 1P	-0.053539	-0.086402	-0.139847	-0.054152	-0.094595	-0.183734
UNP 1P	-0.046402	-0.072970	-0.115340	-0.046240	-0.071848	-0.111500

The results indicate that the VaR level for the NIG and GH are mixed; in other words, the VaR level of the NIG for this data set is not necessarily smaller than the GH as it generally should be. However, for the CVaR, as a better approximation for studying risk exposure, the GH presents more significant absolute expected losses, yet their differences are marginal. Refer to Annex III for graphical representation of the three confidence VaR levels of the NIG (blue line color) and GH (red line color) for the out-of-sample data analysis. Also, Table 15 exposes the computational effort required for both models. GH requires a higher level of computational efforts than NIG, even though the AIC results show no significant difference between them.

Table 15. Computational effort evaluation.

	NIG			GH		
Asset / Period	Log-likehood	AIC	Iterations	Log-likehood	AIC	Iterations
ADBE 1P	5,219.05	-10,430.09	141	5,219.05	-10,428.10	254
AMZN 1P	4,828.44	-9,648.89	139	4,829.33	-9,648.67	340
DIS 1P	6,184.32	-12,360.64	231	6,185.98	-12,361.96	362
NEE 1P	7,057.38	-14,106.76	243	7,060.29	-14,110.58	478
NKE 1P	6,354.37	-12,700.74	147	6,354.50	-12,699.00	304
NVDA 1P	4,382.02	-8,756.03	233	4,385.67	-8,761.34	502
UNH 1P	6,223.06	-12,438.12	155	6,230.74	-12,451.47	502
UNP 1P	6,429.77	-12,851.55	219	6,429.97	-12,849.93	388

4.6 CONCLUSION

This chapter analyzed eight financial assets that present a bubble-type behavior during two periods of analysis: from January 3, 2000, to December 31, 2009 (1P), and from January 4, 2010, to April 29, 2020 (2P). The first period relates to a part of the dotcom bubble and the 2008 financial crisis, while the second period belongs to more than a decade of continuous growth in the major stock indices around the world.

Hence, the candidate distribution to fit the first segment 1P was the NIG, a hyperbolic family member. The results present that it was able to fit all the assets' first period we studied. It is noteworthy that the NIG distribution has multiple properties, such as being well adjusted to heavy tails and close under convolution.

Using statistical tests, we could confirm the NIG manages to fit the stocks of Amazon.com Inc., UnitedHealth Group Incorporated, NVIDIA Corporation, Walt Disney Company, Adobe Inc., NextEra Energy Inc., NIKE Inc. Class B, and Union Pacific Corporation, where they have multiple exuberant episodes from 2000 to 2020.

Furthermore, by comparing the NIG with the GH, we established that the GH differences are just marginal, and it is not always superior to the NIG on empirical data sets. Finally, the NIG displays exceptional performance in the out-of-sample VaR and CVaR with the parameters adjusted from the in-sample data.

5. CHAPTER IV: A MODEL-BASED CLUSTERING APPROACH

5.1 INTRODUCTION

The economic problem is a human issue, not a natural one, at least for those above the subsistence level. In other words, scarcity depends not only on nature capacity but also on human nature (Heilbroner & Milberg, 2008). Ex nihilo nihil fit, meaning that nothing comes from nothing, or from nothing, nothing comes. Therefore, we can argue that the economic problem has an origin, and the financial bubbles arise from some irrational behavior or expose human nature once again.

"Civilizations form, evolve, and sometimes collapse and disappear, leaving behind their remnants in the form of monuments, burial sites, and crafts in the form of art or writings. From these archaeologists decipher details about each civilization" (Mobasher, 2018, p. 314). Hence, we should be financial archaeologists, understanding the complex dynamic and evolution of asset price bubbles.

Since the origin of financial markets, they have shown us that they have the capacity that, in a short period, can go from invigorating to disastrous. The concept of "financial bubble" has become a popular term in the media to conceive an explosion and collapse in financial assets. Nevertheless, we must remember that their behavior is not as simple as it sounds due to its complex nature and evolution.

We usually read the same authors that explain the definition of financial bubbles; however, we got an original reference from literature where Joseph De la Vega (1957), in his book "Confusión de Confusiones" conceptualizes this term as follows: "… they offer for the stocks more than the price of the day (what we call "inflating" the price). They influence the price in this way to sell (short) at the higher figure and thus gain in the end. God with one breath breathed life into Adam, whereas the bears take the life of many people by inflating the price (of the shares) …" (p. 33). A significant part of history where greedy and other human desires connect since many years ago. So, bubbles are not something new slightly; but we still misunderstood them.

Clustering analysis is essential for identifying patterns or specific groups with similarities or dissimilarities. It has a broad academic application because we are always trying, or we tend to develop in our mind some pattern recognition. The purpose of pattern recognition is to assign classes to objects regarding some variables or properties. Thus, the general goal is to distinguish, classify, or generate groups from a known or unknown object division. We are continuously searching for pattern recognition of natural or artificial objects (Beyerer, Richter, and Nagel, 2017). Nonetheless, we have to take one step forward and establish a subtle difference between clustering and classification; in the latter, we also have to define the classes (Gan, Ma, and Wu, 2007). From this short description, we can deduce there is extensive research in different fields of knowledge, and there could be an opportunity cost in financial bubbles.

Consequently, we will conduct a clustering analysis in a modest attempt to fill the literature gap by empirically analyzing the categorization of financial bubbles using the Gaussian finite mixture model. Hence, asset price bubbles categorization is a clustering analysis based on the "bubble size" and "crash size." The first one is the growth percentage from the lowest price (P1) to the highest price (P2), while the latter refers to the collapse percentage from the highest price (P2) to the subsequent lowest price (P3) in a certain period. Subsequently, model-based clustering analysis for asset price bubbles could work as a dynamic tool for the study of ex-ante and ex-post bubble-type behavior.

5.2 GAUSSIAN FINITE MIXTURE MODEL

Model-based clustering or finite mixture model surges as an alternative from heuristic approaches where we consider the data as coming from a probability distribution (Fraley & Raftery, 2002). Finite mixture models assume that a population is a convex combination of a finite number of densities. A random vector X results from a parametric finite mixture distribution if

$$for\ \forall x \subset X \tag{1}$$

We can write the density as follows

$$f(x|\Psi) = \sum_{k=1}^{K} \pi_k f_k(x|\theta_k) \tag{2}$$

where $\pi_k > 0$, and $\sum_{k=1}^{K} \pi_k = 1$ are the mixing coefficients or mixing proportions, $f_1(x|\theta_k), \dots, f_K(x|\theta_K)$ are the component densities, and $\Psi = (\pi, \theta_1, \dots, \theta_K)$ is the parameters' vector with $\pi = (\pi_1, \dots, \pi_K)$. Usually, the component densities are of the same type of distribution, and they tend to be multivariate Gaussian (Browne & McNicholas, 2015). "The idiom 'model-based clustering' is used to connote clustering using mixture models" (Browne & McNicholas, 2015, p. 177).

Banfield and Raftery (1993) proposed a framework for model-based clustering to overcome the limitations clustering analysis has in those years. The limitations of the classification maximum likelihood procedure were the following:

1. The covariance matrices are constant, so it only considers the restrictive model; however, it would be appropriate that covariance matrices could differ between clusters.
2. It allows only for Gaussian distributions.
3. In general, the procedure does not allow noise or data points that do not fit the clusters' main pattern.

Consequently, they developed maximum likelihood criteria for Gaussian clustering to allow some changes in the covariance matrices. They also extended the model to contain Poisson noise and showed a Bayesian approach regarding the selection of the number of clusters.

Fraley and Raftery (1998) exposed a clustering methodology where they applied the Bayesian information criterion (BIC) to compare models that may differ from the number of components and the underlying densities. Moreover, they described how the Expectation-Maximization (EM) algorithm could provide a measure of uncertainty about the resulting clusters. For previous introductory investigations over cluster analysis, we can search for Hartigan (1975), Gordon (1981), McLachlan and Basford (1988), and Kaufman and Rousseeuw (2009).

Fraley and Raftery (1999) developed a software package named MCLUST for cluster analysis applying parameterized Gaussian hierarchical clustering algorithms and the Expectation-Maximization (EM) algorithm for optimization with the possible addition of a Poisson noise term. This method has shown positive results in practical applications such as character recognition, tissue segmentation, minefield, seismic fault detection, identification of textile flaws from images, and astronomical data classification.

Furthermore, Scrucca, Fop, Murphy, and Raftery (2016) updated the famous R package named MCLUST for clustering, classification, and density estimation based on Gaussian finite mixture models. In this paper, they developed "newly available models, dimension reduction for visualization, bootstrap-based inference, implementation of different model selection criteria and initialization strategies for the EM [Expectation-Maximization] algorithm" (p. 290). In the model-based clustering approach, clusters are ellipsoidally centered at the mean vector while the covariance matrix determines other geometric features (volume, shape, and orientation).

We obtained parsimonious parameterizations by means of an eigendecomposition of the covariance matrix Σ_k, where the eigendecomposition is as follows

$$\Sigma_k = \lambda_k D_k A_k D_k^T \quad (3)$$

where λ_k controls the volume of the ellipsoid and is a scalar, A_k is a diagonal matrix specifying the shape of the density contours with $det(A_k) = 1$ and D_k

determines the orientation of the ellipsoid and is an orthogonal matrix (Banfield & Raftery, 1993; Celeux & Govaert, 1995). Therefore, we can specify fourteen models with the MCLUST new version according to different geometric characteristics and in Fig. 26, we show them (Scrucca et al., 2016). The models are EII, VII, EEI, VEI, EVI, VVI, EEE, EVE[6], VEE[7], VVE[8], EEV, VEV, EVV[9], and VVV. See Table 16.

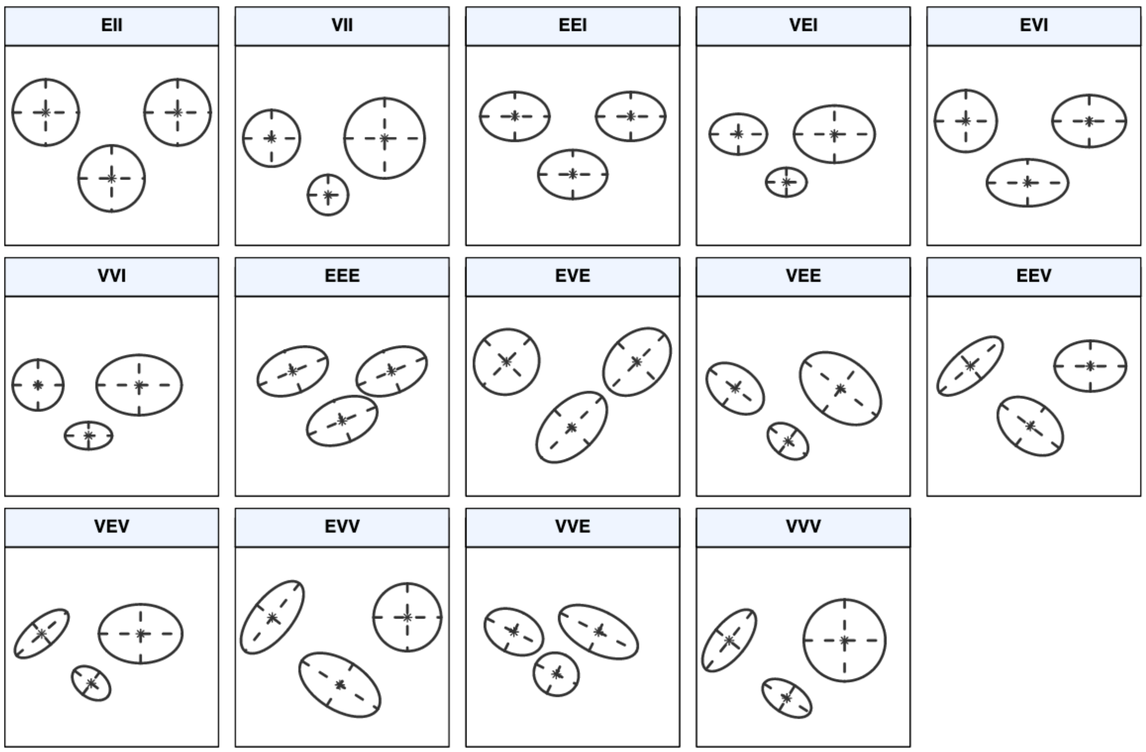

Fig. 26 Ellipses of the fourteen Gaussian models obtained by eigen-decomposition for three groups (Scrucca et al., 2016, p. 292).

[6] This model was not a member of the MCLUST family until new versions of the software.
[7] This model was not a member of the MCLUST family until new versions of the software.
[8] This model was not a member of the MCLUST family until new versions of the software.
[9] This model was not a member of the MCLUST family until new versions of the software.

Table 16. Parameterizations of the covariance matrix for multidimensional data in the MCLUST package (Scrucca et al., 2016).

Model	Σ_k	Distribution	Volume	Shape	Orientation
EII	λI	Spherical	Equal	Equal	—
VII	$\lambda_k I$	Spherical	Variable	Equal	—
EEI	λA	Diagonal	Equal	Equal	Coordinate axes
VEI	$\lambda_k A$	Diagonal	Variable	Equal	Coordinate axes
EVI	λA_k	Diagonal	Equal	Variable	Coordinate axes
VVI	$\lambda_k A_k$	Diagonal	Variable	Variable	Coordinate axes
EEE	λDAD^T	Ellipsoidal	Equal	Equal	Equal
EVE	λDA_kD^T	Ellipsoidal	Equal	Variable	Equal
VEE	$\lambda_k DAD^T$	Ellipsoidal	Variable	Equal	Equal
VVE	$\lambda_k DA_kD^T$	Ellipsoidal	Variable	Variable	Equal
EEV	$\lambda D_kAD_k^T$	Ellipsoidal	Equal	Equal	Variable
VEV	$\lambda_k D_kAD_k^T$	Ellipsoidal	Variable	Equal	Variable
EVV	$\lambda D_kA_kD_k^T$	Ellipsoidal	Equal	Variable	Variable
VVV	$\lambda_k D_kA_kD_k^T$	Ellipsoidal	Variable	Variable	Variable

Furthermore, in many practical applications, the number of mixture components $\boldsymbol{K}$ is unknown, so we use the Bayesian information criterion (BIC) to make our model selection. It has become a famous model selection because other alternatives have not yet proven superior and is a common choice in the Gaussian mixture modeling literature (Browne & McNicholas, 2015).

5.3 MATHEMATICS BEHIND THE GAUSSIAN FINITE MIXTURE MODEL

The Gaussian distribution may not apply in practice, but many analysts are still using it in many applications because the Gaussian distribution gives us an easy mathematical manipulation. The popularity of the Gaussian distribution is due to its mathematical tractability and flexibility for density estimation. Nevertheless, there may be cases where the Gaussian distribution will not fit. In other words, the nature of the sample is “stickily,” not Gaussian. In such situations, we can deal with multiple Gaussian distributions, for example.

The univariate Gaussian distribution,

$$G(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4}$$

where, $\boldsymbol{\mu}$ is the mean and $\boldsymbol{\sigma}$ is the standard deviation ($\boldsymbol{\sigma^2}$ is the variance). And we can extend it to a multivariate Gaussian distribution

$$\mathcal{N}(x|\mu,\Sigma) = \frac{1}{(2\pi|\Sigma|)^{\frac{1}{2}}} exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\} \tag{5}$$

where, $\boldsymbol{\mu}$ is a mean’s vector of dimension $\boldsymbol{x}$, and $\boldsymbol{\Sigma}$ is the covariance.

Now, we can extend $\boldsymbol{\mathcal{N}}$ in order to have multiples of multivariate Gaussian distribution. Hence, we need to estimate the parameters $(\boldsymbol{\Sigma}, \boldsymbol{\mu})$. We could think of applying the usual method of Maximum Likelihood Estimation (MLE); however, we will not arrive at a closed-form.

If we have a mixture of Gaussian as follows

$$p(x) = \sum_{k=1}^{K} \pi_k \, \mathcal{N}(x|\mu_k, \Sigma_k) \quad (6)$$

where $\boldsymbol{K}$ is the number of Gaussian distributions, $\boldsymbol{\pi_k}$ is the mixing coefficient, and as we mentioned above $\boldsymbol{\mathcal{N}(x|\mu_k, \Sigma_k)}$ is the multivariate Gaussian distribution for the $\boldsymbol{k}$th Gaussian.

Additionally, normalization and positivity require that,

$$0 \leq \pi_k \leq 1 \quad (7)$$

$$\sum_{k=1}^{K} \pi_k = 1 \quad (8)$$

It is noteworthy that the mixing coefficient is also known as the weights of the corresponding Gaussian distribution, where the log-likelihood of a mixture of Gaussian is as follows

$$ln\, p(X|\mu, \Sigma, \pi) = \sum_{n=1}^{N} \ln p(x_n) = \sum_{n=1}^{N} ln \left\{ \sum_{k=1}^{K} \pi_k \, \mathcal{N}(x_n|\mu_k, \Sigma_k) \right\} \quad (9)$$

where $\boldsymbol{X}$ represents $\boldsymbol{n}$ numbers of $\boldsymbol{x}$, $\boldsymbol{x_n}$ refers to a particular number of the sample, $\boldsymbol{N}$ is the total number of samples. As we specified before, the MLE does not work as there is no closed-form solution.

Consequently, we can calculate the parameters using the Expectation-Maximization (EM) technique, which is an iterative method for optimization.

We can consider mixing coefficients as prior probabilities for the components. For a given value $\boldsymbol{x}$, we can calculate the corresponding posterior probabilities, which are also called responsibilities, where they are like some latent variables.

From Bayes rule, we have

$$\gamma_k(x) = p(k|x) = \frac{p(k)p(x|k)}{p(x)} \tag{10}$$

$$= \frac{\pi_k \mathcal{N}(x|\mu_k, \Sigma_k)}{\sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)} \tag{11}$$

$$\pi_k = \frac{N_k}{N} \tag{12}$$

where $\boldsymbol{k}$ is a latent variable under Bayes definition, and $\boldsymbol{\pi_k}$ is the mixing coefficient with $\boldsymbol{N_k}$ as the number of samples for a particular class divided by $\boldsymbol{N}$ as the total number of samples.

So, we are taking the conditional probability to be in the numerator with the mixing coefficients and the mixture of Gaussians, and in the unconditional probability in the denominator with the sum of the mixture of Gaussians. Therefore, we can understand $\boldsymbol{N_k}$ as the number of points that are assigned to a specific cluster $\boldsymbol{k}$. However, the number of samples that will belong to a particular cluster $\boldsymbol{k}$ is not previously known, and we will obtain it from our calculations.

Based on the information mentioned above, the EM algorithm will help us to establish the parameters through the next steps:

1. Estimation step: for some given parameters, we can calculate the expected values of the latent variable.
2. Maximization step: using the MLE method, we will update the parameters based on the calculated latent variable.

Our objective is to maximize the likelihood function given a Gaussian mixture model regarding the parameters containing the means and covariances of the components and the mixing coefficients.

First, we have to initialize the means $\boldsymbol{\mu_k}$, covariances $\boldsymbol{\Sigma_k}$, and mixing coefficients $\boldsymbol{\pi_k}$. Next, we have to evaluate the initial value of the log-likelihood. Second, evaluate the responsibilities with the actual parameters, referring to the estimation step of the EM algorithm

$$\gamma_k(x) = \frac{\pi_k \mathcal{N}(x|\mu_k, \Sigma_k)}{\sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)} \tag{13}$$

Third, for the maximization step, we re-estimate the parameters using the current responsibilities as follows

$$\mu_k = \frac{\sum_{n=1}^{N} \gamma_k(x_n) x_n}{\sum_{n=1}^{N} \gamma_k(x_n)} \tag{14}$$

$$\sum_k = \frac{\sum_{n=1}^{N} \gamma_k(x_n)(x_n - \mu_k)(x_n - \mu_k)^T}{\sum_{n=1}^{N} \gamma_k(x_n)} \tag{15}$$

$$\pi_k = \frac{1}{N} \sum_{n=1}^{N} \gamma_k(x_n) \tag{16}$$

Finally, we calculate the log-likelihood

$$ln\, p(X|\mu, \Sigma, \pi) = \sum_{n=1}^{N} ln \left\{ \sum_{k=1}^{K} \pi_k\, \mathcal{N}(x_n|\mu_k, \Sigma_k) \right\} \tag{17}$$

Besides, if there is no converge, the computation repeats the steps from the Expectation-Maximization algorithm until the parameters or the log-likelihood do not change after successive iterations, depending on the convergence criteria.

5.4 ASSET PRICE BUBBLE CATEGORIZATION

Asset price bubble analysis is fundamental due to investors' irrational euphoria in an environment where prices keep rising and seem not to yield to a possible fall. Nevertheless, when they plunge, it is like a waterfall without an immediate end. Past financial collisions can give us an explanation of how markets never drive concerning their price action.

We took the daily closing prices in US dollars from Bloomberg of the financial bubbles studied in Zhang, Zhang, and Sornette (2016) to carry their categorization. See Table 17.

Table 17. We present the list of sixteen historical bubbles analyzed in Zhang et al. (2016).

Number	Asset and year of crash	Period
1	S&P 500 1987	02/01/1984 – 13/11/1987
2	S&P 500 2007[10]	01/01/2004 – 31/12/2009
3	DJIA 1929	02/01/1926 – 31/12/1930
4	Nasdaq Composite Index 2000	01/01/1993 – 31/12/2002
5	Chile 1991 / 1994	01/10/1987 – 01/12/2000
6	Venezuela 1997	03/01/1994 – 30/12/1999
7	Indonesia 1994 / 1997	03/01/1990 – 30/12/1999
8	Malaysia 1994	01/01/1991 – 29/12/1995
9	Thailand 1994	01/01/1990 – 30/12/1994
10	Hong Kong 1987 / 1994 / 1997	02/01/1980 – 31/12/1999
11	Hong Kong 2007	03/01/2000 – 10/04/2015
12	Sugar price	01/01/2002 – 31/12/2013
13	Brent Oil 2008	01/01/1990 – 16/04/2015
14	SSEC 2007 / 2009	01/01/2004 – 31/12/ 2014
15	SZSC 2007 / 2009	01/01/2004 – 31/12/ 2014
16	SSEC 2015	23/02/2011 – 12/05/2015

We do not include 5, 6, 12, and 13 because we do not obtain the required period's complete time series. While for numbers 7, 10, 14, and 15, it is important to mention they comprise more than one bubble. We also extract the S&P 500 index close prices of 2020 (S&P 500 2020)[11] as an extension of the above Table, which we will apply in an analysis of price fall scenarios.

[10] We consider the initial date from January 2, 2004, since January 1, 2004, is holiday.
[11] Until March 20, 2020.

Therefore, to categorize asset price bubbles, we calculated the “bubble size” and “crash size” of the exuberance in the financial assets of Table 17. Additionally, in Table 18 and Fig. 27, we show the mentioned analysis and graphic representation.

Table 18. Bubble size and crash size analysis.

Asset	P1 date	P1	P2 date	P2	P3 date	P3	Bubble size	Crash size
S&P 500 1987	24/07/1984	147.82	25/08/1987	336.77	19/10/1987	224.84	127.82%	33.24%
S&P 500 2007	12/08/2004	1,063.23	09/10/2007	1,565.15	09/03/2009	676.53	47.21%	56.78%
DJIA 1929	30/03/1926	135.2	03/09/1929	381.17	16/12/1930	157.51	181.93%	58.68%
NASDAQ 2000	26/04/1993	645.87	10/03/2000	5,048.62	09/10/2002	1,114.11	681.68%	77.93%
INDONESIA 1994	30/10/1991	224.706	05/01/1994	612.89	19/04/1995	414.21	172.75%	32.42%
INDONESIA 1997	19/04/1995	414.209	08/07/1997	740.83	21/09/1998	256.83	78.85%	65.33%
MALAYSIA	16/01/1991	470.4	05/01/1994	1,314.46	24/01/1995	840.87	179.43%	36.03%
THAILAND	30/11/1990	544.3	04/01/1994	1,753.73	04/04/1994	1,196.18	222.20%	31.79%
HONG KONG 1987	02/12/1982	676.3	01/10/1987	3,949.73	07/12/1987	1,894.94	484.02%	52.02%
HONG KONG 1994	07/12/1987	1,894.94	04/01/1994	12,201.09	23/01/1995	6,967.93	543.88%	42.89%
HONG KONG 1997	23/01/1995	6,967.93	07/08/1997	16,673.27	13/08/1998	6,660.42	139.29%	60.05%
HONG KONG 2007	25/04/2003	8,409.01	30/10/2007	31,638.22	27/10/2008	11,015.84	276.24%	65.18%
SSEC 2007	11/07/2005	1,011.499	16/10/2007	6,092.06	04/11/2008	1,706.70	502.28%	71.98%
SSEC 2009	04/11/2008	1,706.703	04/08/2009	3,471.44	27/06/2013	1,950.01	103.40%	43.83%
SSEC 2015	27/06/2013	1,950.012	27/04/2015	4,527.40	26/08/2015	2,927.29	132.17%	35.34%
SZEC 2007	15/11/2005	2,622.03	31/10/2007	19,531.15	04/11/2008	5,668.81	644.89%	70.98%
SZEC 2009	04/11/2008	5,668.81	07/12/2009	14,051.52	20/03/2014	6,998.19	147.87%	50.20%

Fig. 27 Bubble size and crash size scatter plot.

We use the MCLUST package to categorize the "bubble size" and "crash size" because it is recognized as a gold standard approach for model-based clustering. Thus, we propose the Gaussian mixture model to fill the literature gap in financial bubbles and use it as the primary model for further research. Commonly, the letter "G" denotes the number of clusters, so we continue with this notation while selecting the best model according to the Bayesian information criterion (BIC).

The current financial collapse that occurred during the COVID-19 pandemic break out shows how sensitive financial markets are. They arrived at a stage of more than a decade[12] of continuous growth until this disease appeared, and markets tumbled down. These falls motivated us to analyze three possible scenarios in the S&P 500 index and explore the variations in the new scenarios and their impact on clustering analysis. The "bubble size" and "crash size" scenarios are as follows:

[12] The date of this investigation is until May 20, 2020.

Table 19. Possible crash scenarios in the S&P 500 index.

Asset	P1 date	P1	P2 date	P2	P3 date	P3	Bubble size	Crash size
S&P 500 2020 (S1)	02/07/2010	1,022.58	19/02/2020	3,386.15	#N/A	2,271.72	231.14%	32.91%
S&P 500 2020 (S2)	02/07/2010	1,022.58	19/02/2020	3,386.15	#N/A	1,565.15	231.14%	53.78%
S&P 500 2020 (S3)	02/07/2010	1,022.58	19/02/2020	3,386.15	#N/A	676.53	231.14%	80.02%

It is noteworthy that the P3 date column is not filled because we do not consider the date's forecast by which prices would reach these scenarios. S1, S2, and S3 denote our fall price scenarios, while P1 and P2 in all cases are the same numbers because we set the lowest and highest price after the 2008 crisis. Moreover, P3 of the first scenario (S1) relates to a price resistance between 2015 and 2016. For the price P3 in S2 and S3, they appear as the prices for P2 and P3 of the S&P 500 2007 asset price bubble in Table 19.

We implemented the MCLUST, and the BIC selected a VEV five model component (G=5). Fig. 28 shows the model selection regarding the BIC and the optimal number of components. Fig. 29 displays the model-based clustering analysis for the "bubble size" and "crash size." Fig. 30 illustrates the uncertainty plot, where more significant observations in the graph indicate the major uncertainty of the data belonging to a particular cluster. Table 20 shows a summary of the results and parameters obtained.

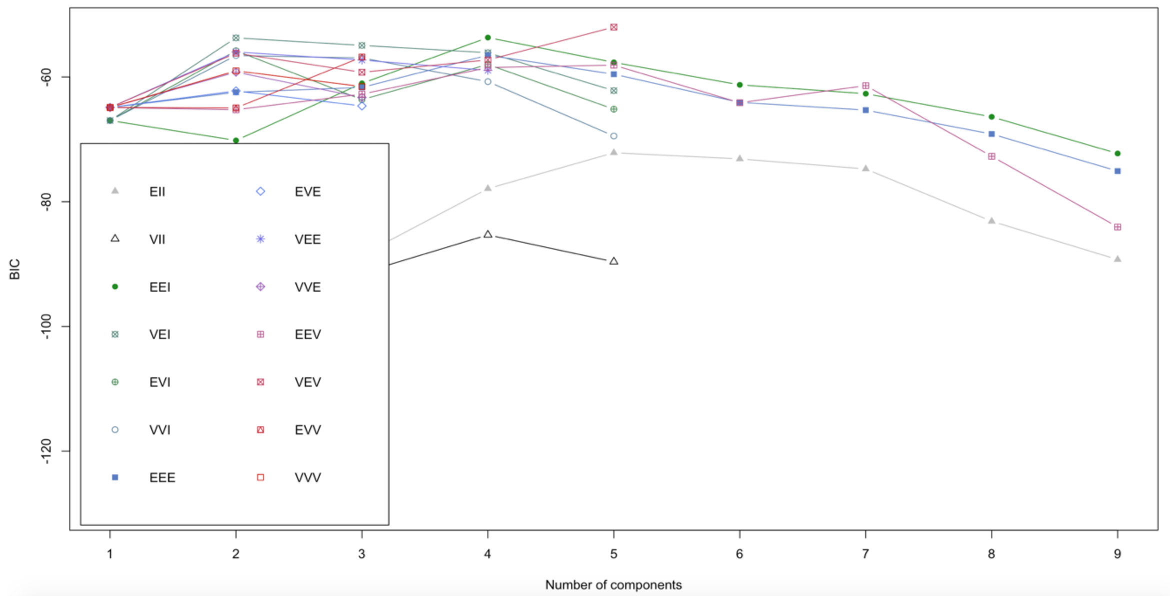

Fig. 28 BIC plot for models fitted to financial bubbles categorization.

The summary of the selected model is as follows:

Table 20. Gaussian finite mixture model fitted by EM algorithm.

-- Gaussian finite mixture model fitted by EM algorithm --
Mclust VEV (ellipsoidal, equal shape) model with 5 components: log-likelihood n df BIC ICL 9.414961 17 25 -52.00041 -52.00425 Clustering table: 1 2 3 4 5 5 5 3 2 2
Mixing probabilities: 1 2 3 4 5 0.2941176 0.2941330 0.1764580 0.1176471 0.1176443
Means:

	[,1]	[,2]	[,3]	[,4]	[,5]
BUBBLE SIZE	1.6688	1.4475186	6.0958013	5.1395	1.2564947
CRASH SIZE	0.3374	0.6122145	0.7363163	0.4745	0.4699992

Variances:

[,,1]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.199317961	-0.0040010549
CRASH SIZE	-0.004001055	0.0002372953

[,,2]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.94438105	0.01741074
CRASH SIZE	0.01741074	0.00106467

[,,3]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.70986632	0.0154336949
CRASH SIZE	0.01543369	0.0008947103

[,,4]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.044850940.	-0.006808263
CRASH SIZE	-0.006808263	0.001070420

[,,5]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.024753528	0.0035571985
CRASH SIZE	0.003557198	0.0005314794

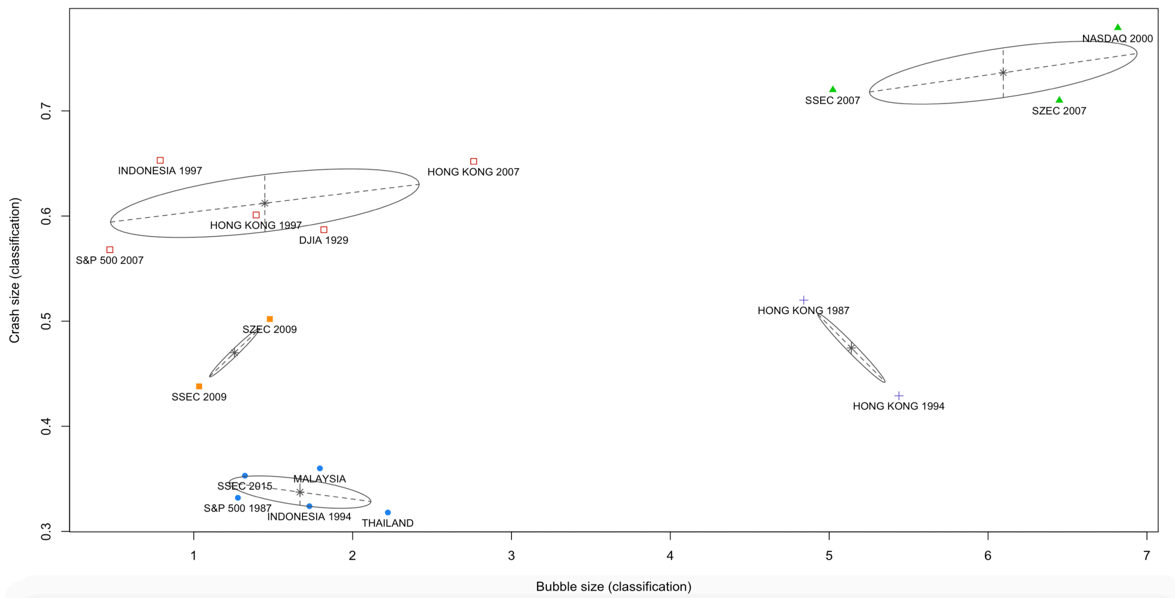

Fig. 29 Estimated cluster plot classification.

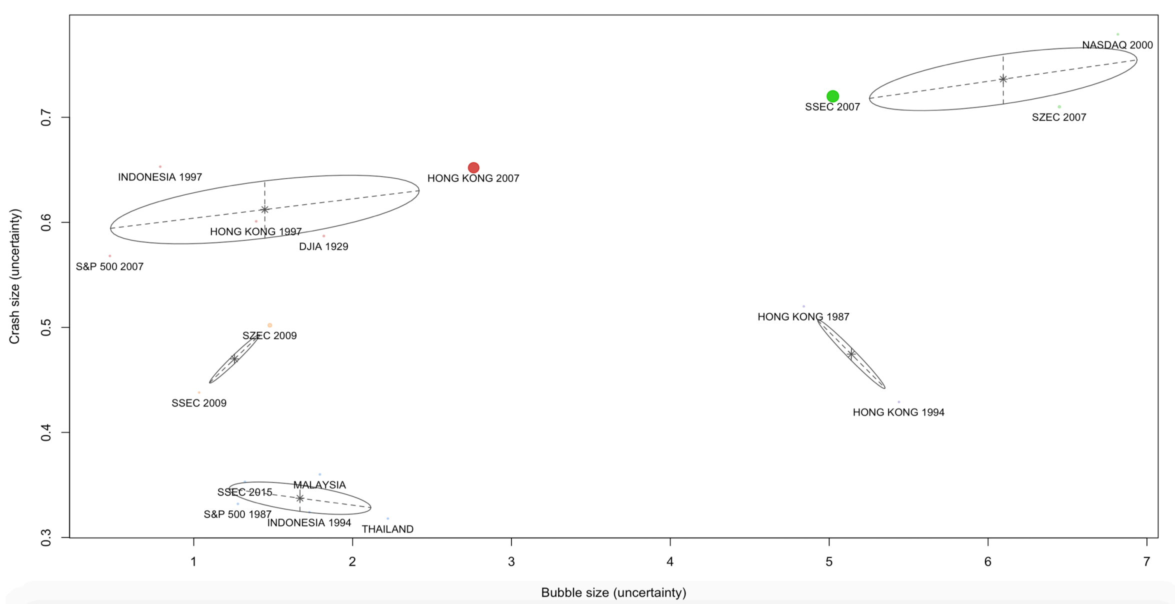

Fig. 30 Estimated cluster plot uncertainty.

In Fig. 29 we observe the five clusters as follows: 1) NASDAQ 2000, SZEC 2007, SSEC 2007; 2) HONG KONG 2007, INDONESIA 1997, HONG KONG 1997, DJIA 1929, S&P 500 2007; 3) SZEC 2009, SSEC 2009; 4) HONG KONG 1987, HONG KONG 1994; and 5) SSEC 2015, MALAYSIA, S&P 500 1987, INDONESIA

1994, THAILAND. While in Fig. 30 HONG KONG 2007 and SSEC 2007 show more uncertainty respect to their clusters.

Secondly, for the S&P 500 2020 (S1) first fall scenario with a 32.91% crash, the BIC chose a five-component VEV model. Fig. 31 sets the optimum number of clusters, Fig. 32, and Fig. 33 exposes the clustering analysis and the uncertainty plot, respectively. Table 21 shows a summary of the results and parameters obtained.

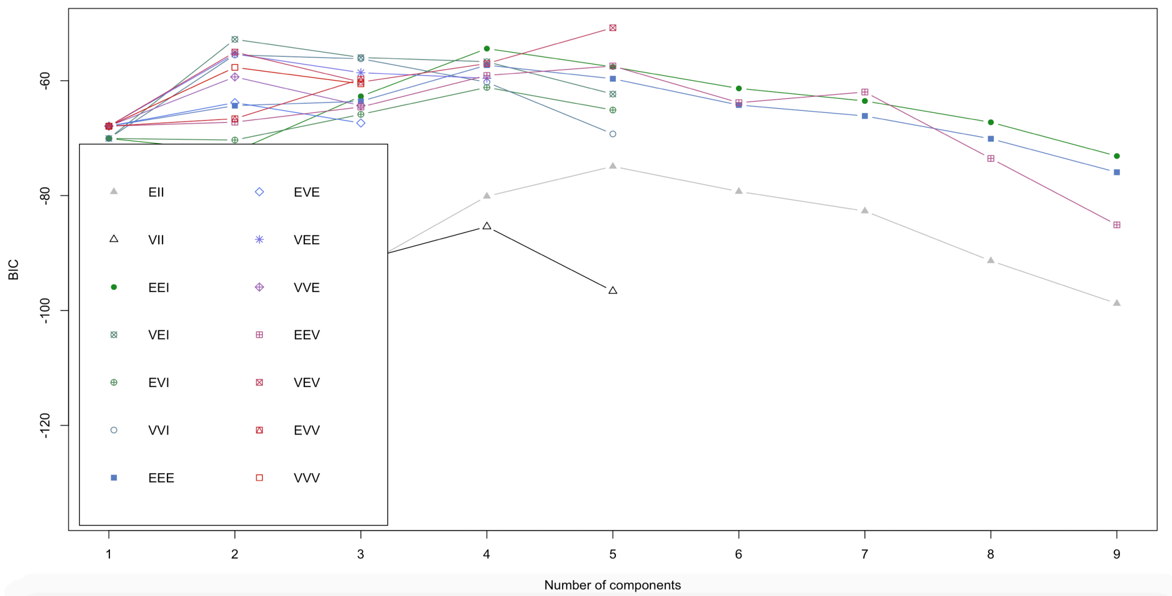

Fig. 31 BIC plot for models fitted to financial bubbles categorization.

The summary of the selected model is as follows:

Table 21. Gaussian finite mixture model fitted by EM algorithm.

--   Gaussian finite mixture model fitted by EM algorithm   --
Mclust VEV (ellipsoidal, equal shape) model with 5 components:

log-likelihood n df BIC ICL

10.73452 18 25 -50.79026 -50.79632

Clustering table:

1 2 3 4 5

6 5 3 2 2

Mixing probabilities:

1 2 3 4 5

0.3333333 0.2777934 0.1666527 0.1111111 0.1111094

Means:

[,1] [,2] [,3] [,4] [,5]

BUBBLE SIZE 1.775833 1.4477807 6.0954219 5.1395 1.2564966

CRASH SIZE 0.336000 0.6122226 0.7363049 0.4745 0.4699995

Variances:

[,,1]

BUBBLE SIZE CRASH SIZE

BUBBLE SIZE 0.211546478 -0.0037056383

CRASH SIZE -0.003705638 0.0002118809

[,,2]

BUBBLE SIZE CRASH SIZE

BUBBLE SIZE 1.02752239 0.018967391

CRASH SIZE 0.01896739 0.001064034

[,,3]

BUBBLE SIZE CRASH SIZE

BUBBLE SIZE 0.76533186 0.0166514803

CRASH SIZE 0.01665148 0.0008941728

[,,4]

BUBBLE SIZE CRASH SIZE

BUBBLE SIZE 0.044850844 -0.006808894

CRASH SIZE -0.006808894 0.001066265

[,,5]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.024753481	0.0035575283
CRASH SIZE	0.003557528	0.0005291864

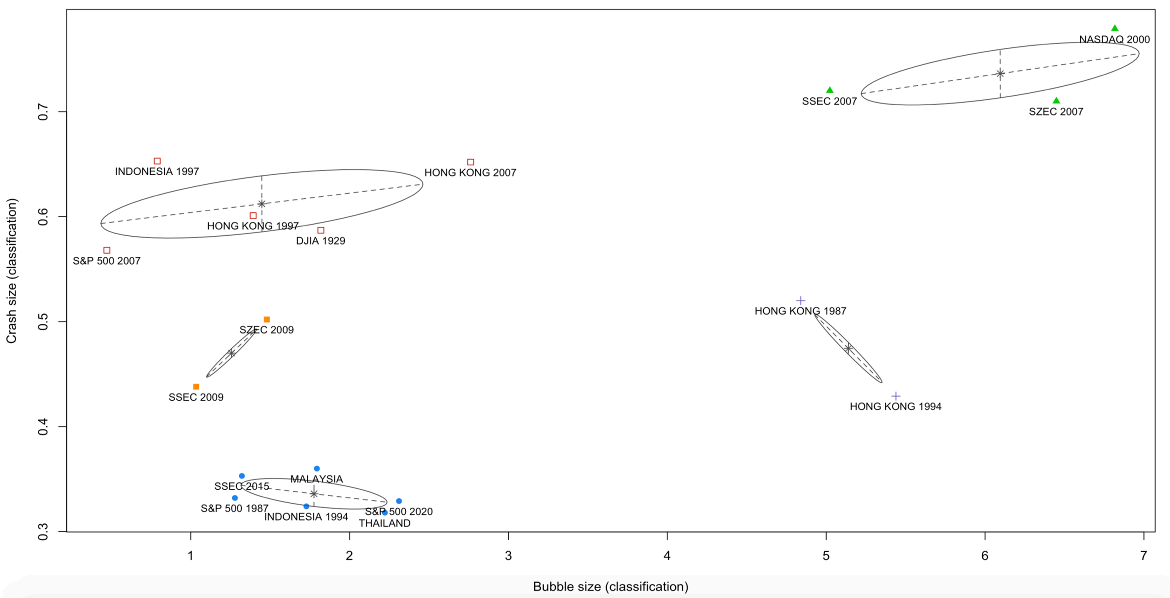

Fig. 32 Estimated cluster plot classification.

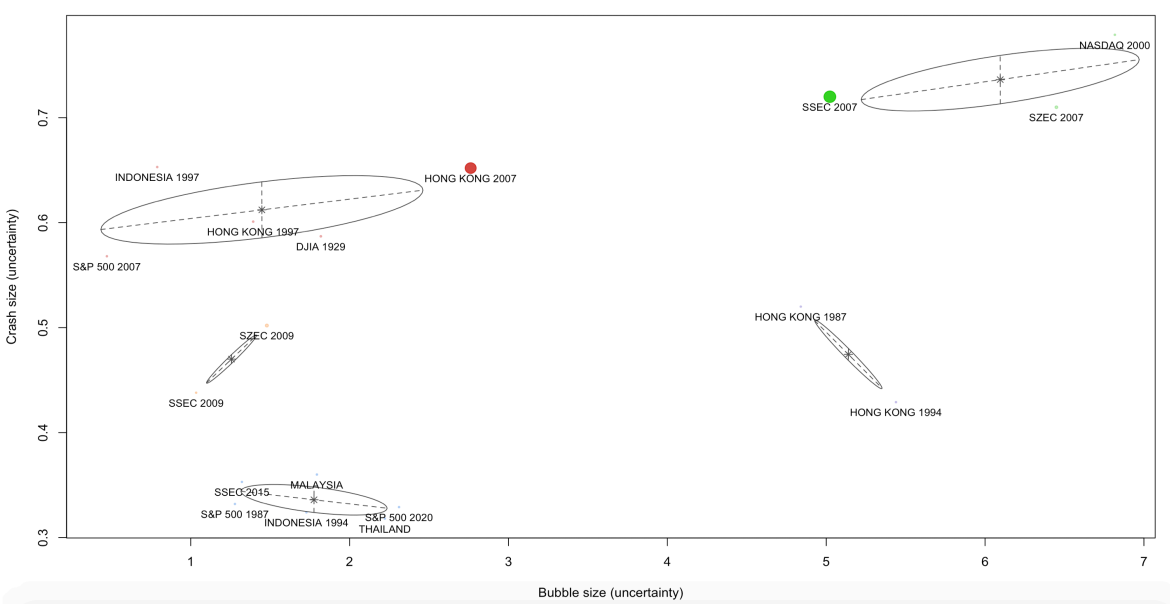

Fig. 33 Estimated cluster plot uncertainty.

Interestingly, the price drop scenario of 32.91% of the S&P 500 2020 (S1) was near to collapse until March 20, 2020, of 31.93%. This new observation (S1) would belong to the fifth cluster mentioned above.

Thirdly, for the S&P 500 2020 (S2) with a price decline scenario of 53.78%, which corresponds to the price (P2) of the S&P 500 2007. The MCLUST adjusted an EEI two model component (G=2) regarding the Bayesian information criterion. Fig. 34 shows the optimum number of clusters; Fig. 35 establishes the Gaussian mixture model; Fig. 36 reveals the uncertainty plot, and Table 22 resumes the results of the model.

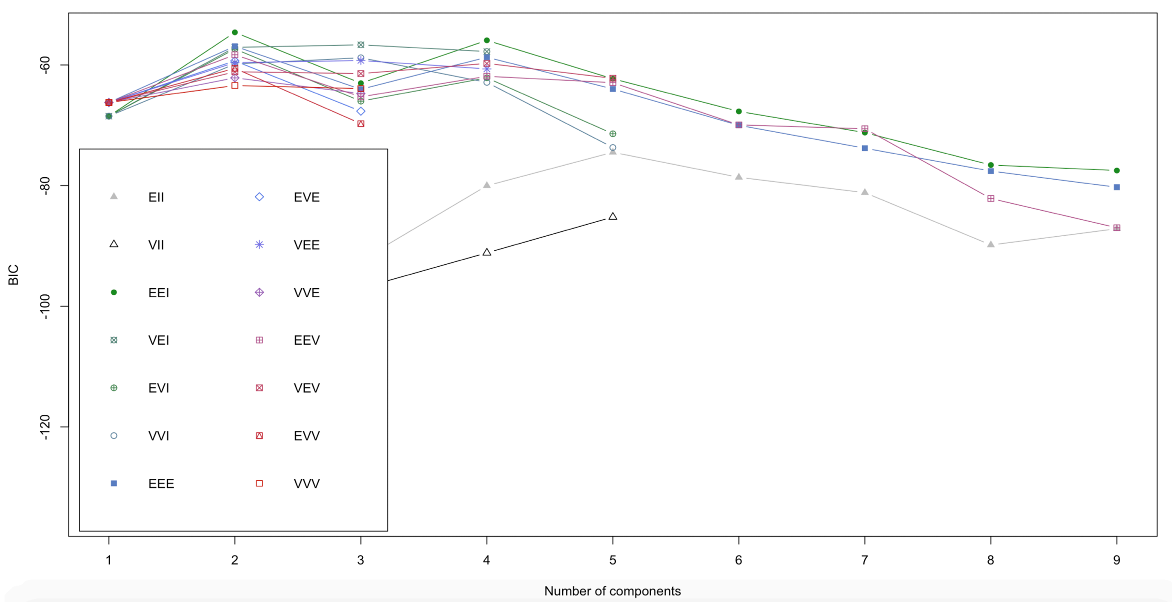

Fig. 34 BIC plot for models fitted to financial bubbles categorization.

The summary of the selected model is as follows:

Table 22. Gaussian finite mixture model fitted by EM algorithm.

```
----------------------------------------------------
Gaussian finite mixture model fitted by EM algorithm
----------------------------------------------------

Mclust EEI (diagonal, equal volume and shape) model with 2 components:

log-likelihood  n   df     BIC         ICL
     -17.18094  18   7   -54.59449   -54.59506

Clustering table:
1   2
13  5

Mixing probabilities:
1           2
0.7222116   0.2777884

Means:
                  [,1]        [,2]
BUBBLE SIZE   1.5694513  5.7134678
CRASH SIZE    0.4789201  0.6316019

Variances:
[,,1]
              BUBBLE SIZE   CRASH SIZE
BUBBLE SIZE   0.438746      0.00000000
CRASH SIZE    0.000000      0.01617437
[,,2]
              BUBBLE SIZE   CRASH SIZE
BUBBLE SIZE   0.438746      0.00000000
CRASH SIZE    0.000000      0.01617437
```

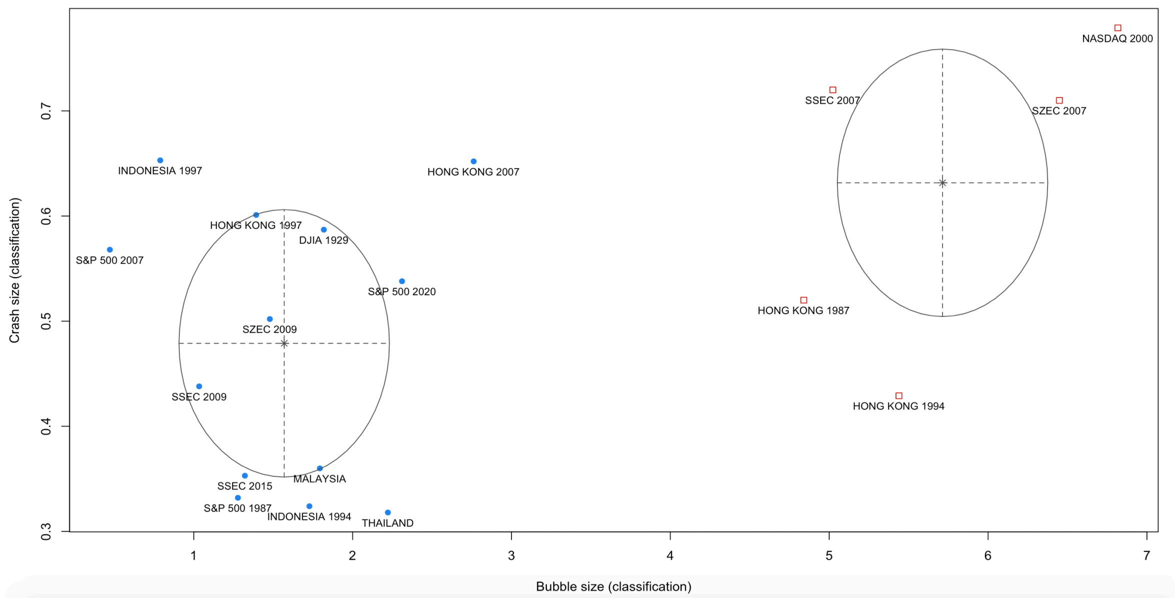

Fig. 35 Estimated cluster plot classification.

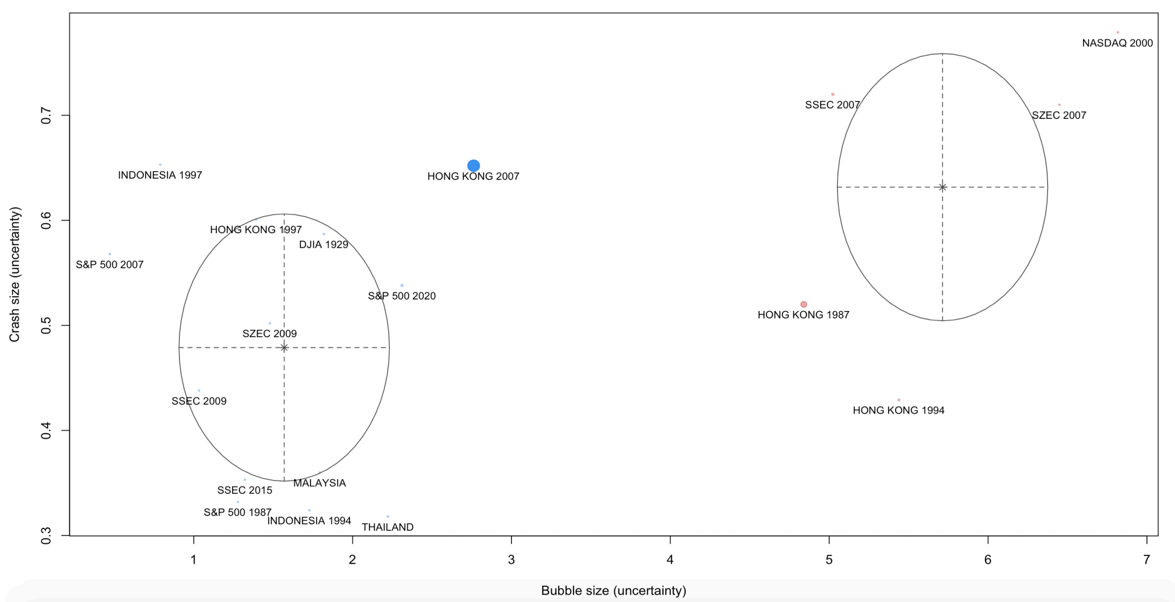

Fig. 36 Estimated cluster plot uncertainty.

The S&P 500 2020 (S2) Gaussian model-based clustering reveals completely different results concerning to above models. The cluster separation is as follows: 1) INDONESIA 1997, HONG KONG 2007, S&P 500 2007, HONG KONG 1997, DJIA 1929, S&P 500 2020, SZEC 2009, SSEC 2009, MALAYSIA, SSEC 2015, S&P 500

1987, INDONESIA 1994, THAILAND; and 2) NASDAQ 2000, SZEC 2007, SSEC 2007, HONG KONG 1987, HONG KONG 1994. On the other hand, HONG KONG 2007 presents the highest uncertainty.

Finally, for the worst-case scenario indicating a crash size of 80.02%, we obtained a VEI model with two components for the S&P 500 2020 (S3). Fig. 37 shows BIC's model selection; Table 23 specifies the model results; Fig. 38 exposes the Gaussian finite mixture model and Fig. 39 represents uncertainty.

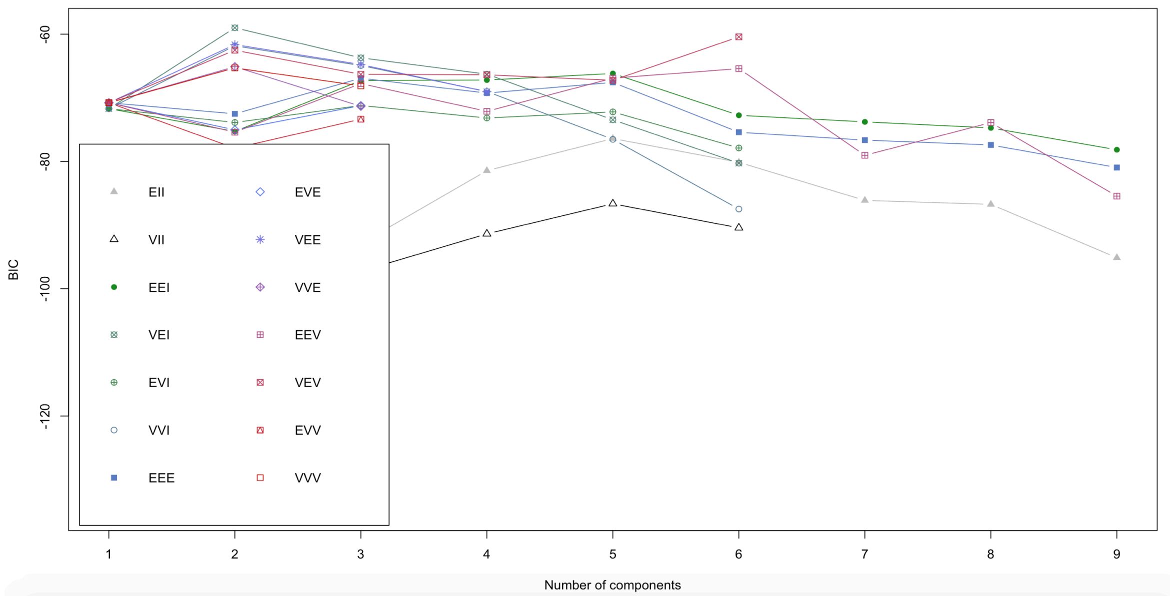

Fig. 37 BIC plot for models fitted to financial bubbles categorization.

The summary of the selected model is as follows:

Table 23. Gaussian finite mixture model fitted by EM algorithm.

--   Gaussian finite mixture model fitted by EM algorithm   --
Mclust VEI (diagonal, equal shape) model with 2 components: log-likelihood n df BIC ICL

-17.93782 18 8 -58.99861 -59.08553

Clustering table:

1 2

5 13

Mixing probabilities:

1 2

0.2753862 0.7246138

Means:

	[,1]	[,2]
BUBBLE SIZE	1.6676928	3.1207680
CRASH SIZE	0.3373952	0.6113255

Variances:

[,,1]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	0.1093337	0.00000000
CRASH SIZE	0.000000	0.0002941306

[,,2]

	BUBBLE SIZE	CRASH SIZE
BUBBLE SIZE	4.903578	0.00000000
CRASH SIZE	0.000000	0.01319165

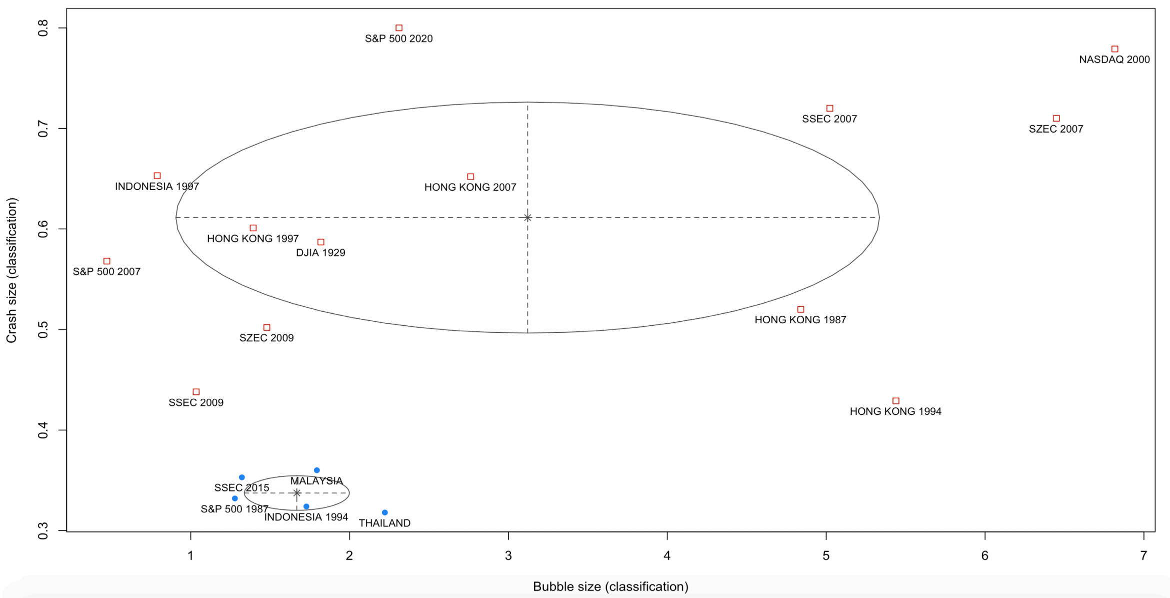

Fig. 38 Estimated cluster plot classification.

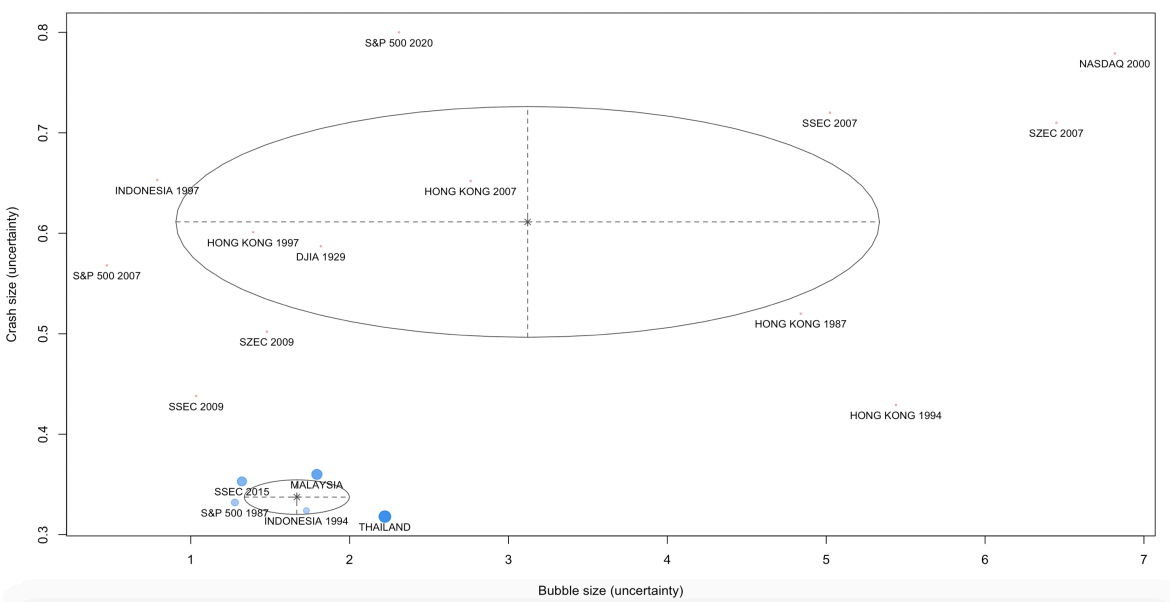

Fig. 39 Estimated cluster plot uncertainty.

We obtained different results from the models fitted by the Expectation-Maximization algorithm, where our results summarized in the next points:

1. A VEV five-component model for the financial bubbles analyzed in Zhang et al. (2016) and our first price drop scenario (S1).
2. An EEI two-component model for the second price scenario (S2).
3. A VEI model with two components for the third price scenario (S3).

We should deduce the Gaussian mixture model is sensitive to the initialization of the EM algorithm. Nevertheless, in the data set from Zhang et al. (2016), we do not obtain different results between running the default parameters of the MCLUST package and a proposed of a better initialization analyzed in Scrucca and Raftery (2015). Therefore, we continued with the default model for the subsequent categorizations. Furthermore, we recognized there are missing crashes like the Tulip Mania, the South Sea Company, Bitcoin crash, to name a few; nevertheless, we wanted to propose categorizing some financial bubbles according to observable variables: "bubble size" and "crash size." We knew that the period studied in our research would be relevant, so we decided to take the initial date and end date from Zhang et al. (2016) as an already published investigation.

5.5 CONCLUSION

In conclusion, we developed a categorization of asset price bubbles applying the Gaussian model-based clustering, where we analyzed four data sets: previous historical crashes and three possible collapse scenarios. We deduced from the previous results the analysis of financial bubbles is and will be a dynamic study because human desires do not stop from one day to another. Humanity and financial bubbles are in continuous systemic instability. It is noteworthy that a useful variable we can incorporate for future research could be market capitalization due to its importance regarding the market size.

6. CHAPTER V: FINAL REMARKS

6.1 CHAPTER II

Financial bubbles represent a failure in human rationality that spreads through the markets and can affect the real economy. It would be fair to assume that time leads to greater wisdom, and such events are less likely to occur. However, the information era has just spread the data and power to many people ignoring past events. Such information flow may be why the increase in volatility and a subsequent irrational behavior with loss-averse agents driven by greed seek to obtain more significant profit regarding the risk or even fundamental analysis. For such, Chapter II presents a monitoring technique that can detect atypical increases in price, leading to the possibility of bubble episodes.

In previous chapters and throughout price asset bubbles literature, the definition of a financial bubble does not converge to a single criterion; however, Sohn and Sornette (2017) proposed two divisions. On the one hand, we have a conditional and comparative definition, and on the other hand, we have the statistical and structural definition. These characterizations are the advanced definitions, not the standard one. We should not forget that the working horse of the bubble type behavior world is the rational model, which assumes that there can be rational deviations from an asset's intrinsic value.

The conditional definition established a relationship between the asset's intrinsic value and the economic agents' information. Therefore, we can understand a financial bubble as a situation where the existing information does not justify the actual price. In contrast, the comparative definition compares the fundamental value with the asset price. The main issue originates in the conceptualization of what we can consider as a fundamental value.

In contrast, the statistical definition contemplates price trajectory, trading volume, and time horizon with no relation to the theoretical background. Bubble's behavior is typically associated with a dramatic increase and a following hard decline. The structural definition explains a specific condition of a particular asset. In other words, this point of view studies the transition of the behavior of the asset. So, we could identify the expansionary and recessionary phases.

We can relate the structural point of view with the regime-switching tests. For example, Homm and Breitung (2012) focus on bubble detection in univariate time series. They took as a benchmark the proposal of Phillips, Wu, and Yu (2011) proposal and compared it to various structural break tests. Their results established that a Chow type break test exhibits the highest power and is the most reliable in finite samples. However, the comparison took the simple test, not the generalized sup Augmented Dickey-Fuller (GSADF) test. The contrast between the latter model and multiple structural breaks (changes) tests could be interesting for further investigations.

It is essential to notice the potential, as well as the constraints of the model. For one side, the GSADF methodology provides a definition and a standardized methodology to detect the bubble that has been born or burst. Furthermore, the generalized proposition with the corrections provided by the wild bootstrap simulation gives a robust technique that combines the ad hoc computation of a window and a recursive technique of a rolling window of extended width to obtain the simulation of the null hypothesis distribution. This simulation procedure with the bootstrap technique benefits from using the data instead of a generic distribution that converges in the limit and improves the critical values' estimation.

The way it works is by selecting a subsample of the data from which the first iterations would be computed. Next, one additional observation is added to the subsample, and the procedure is then repeated. These iterations continue until the last observation (the most recent one) is reached. However, with this procedure, the original model is obtained. The generalized one denoted as GSADF is constructed from the moment the first model is finished. The second step for such a methodology is to move the original window one observation ahead while maintaining the same length. It is equivalent to remove the first observation of the dataset and perform the original model. Once more, these iterations continue until the rolling window reaches the last observation. The sum of the GSADF model uses the type of rolling windows, which increases length for the first part and one of the same lengths for the generalized procedure. Suppose all the aspects are taken into account. In that case, the GSADF methodology uses both a rolling window and a recursive computation to compute the critical values either by Monte Carlo or Bootstrap simulations.

However, some of these benefits also play against the model. In the first place, it is crucial to notice that as with any rolling window procedure, the results would be highly sensitive to changes in the width of such. This document uses the recommendations made by the author to compute the value. However, in a greater perspective, there may be the case that such value over- or under-estimate the unit root test and miss the episodes. The second issue with the model is the computational complexity, as the number of observations increases, the time to perform the simulations. This time increases exponentially with the GSADF methodology and the number of simulations for each period, making the practical and real-time usage dependent on the CPU speed. An extension of these is the frequency of the data. Most of the papers published under the methodology used weekly data, so the historical period analyzed by the broadest possible and remain computationally feasible. The most practical usage would be to use the model with intraday data, but such a quantity of information would make the model unworkable. This topic also could extend to a definition issue, as the bubbles may be considered as a temporary price deviation from fundamental value, then the length of the bubble lifespan is unattended; for daily data, the minimal length is one day, but such atypical value could be the response of technical failures in data recollection. This problem is particularly plausible for intraday data, where jumps are witness with relative frequency.

One final disadvantage of the SADF and GSADF test is that both models do not identify negative bubbles. The concept of negative bubble refers to “as the mirror (but not necessarily exactly symmetric) image of standard financial bubbles, in which positive feedback mechanisms may lead to transient accelerating price falls… The price fall occurring during a transient negative bubble can be interpreted as an effective random down payment that rational agents accept to pay in the hope of profiting from the expected occurrence of a possible rally" (Yan, Woodard, and Sornette, p. 1361, 2012). The negative conceptualization of the financial bubble is related to the log-periodic power law (LPPL) model (Johansen and Sornette, 1999; Jiang, Zhou, Sornette, Woodard, Bastiaensen and Cauwels, 2010), which has in common with the SADF and GSADF tests a view of the financial bubbles that differ from the construction of the fundamental value. Phillips' models consider a transition from a stationary process to a mildly explosive process, while Sornette's model identifies market bubbles as a super-exponential price process. Therefore, the formation of financial bubbles concentrates primarily on the expectation of future earnings rather than present economic value, looking at the explosive behavior as a deviation of a non-sustainable regime. Furthermore, the lecturer could find interesting another kind of models such as the phenomenological Langevin equation model, behavioral models, agent-based models, experimental tests on bubbles, variance bounds tests, West's two-step tests, integration/cointegration based tests, intrinsic bubbles, bubble as an unobserved variable, and algorithm and data-driven prediction model of financial times series (Zhang and Wu, 2018).

Finally, Phillips et al. (2011, 2015) methodologies provide a consistent procedure to define and identify bubble episodes. Nevertheless, the general procedure corresponds with a monitoring rather than a practical technique for trading advantage. This chapter has proven that the episodes have increased in the last years for most markets. However, this result provides evidence for further research regarding the motivation of such behavior. This model corresponds with a data-driven and atheoretical perspective that gives insights for other types of analysis.

6.2 CHAPTER III

This chapter deals with some of the stylized facts described by Cont (2001). This analysis provides evidence that financial prices do not follow a typical Exponential Brownian Motion, so using a normal distribution would lead to a misleading result of the risk of explosive weather behavior or a bubble burst. Such behaviors are associated with high and low return values with an almost cero probability if considered under the normal distribution. As exposed in the previous chapter, these episodes are becoming more usual with time, so this assumption does not hold empirical data. The relevance of this proposal relies mainly on derivative valuation and risk management areas. The reason is that both financial branches use the normality assumption (with the Brownian Motion model) to deploy theoretical or simulation scenarios to value derivatives or to determine the VaR of a specific position. As it occurs in any scientific procedure, if the assumptions are wrong, then the results will undoubtedly have an error that could be catastrophic for financial institutions or any agents using them.

This chapter then proposes a solution for this erroneous assumption. As an alternative, the NIG distribution is proposed as a better fit for empirical data. The methodology to test this new assumption relies on the statistical test performed in two main steps: the normality test and the goodness of fit. As Thode (2002, p. 99) expressed, the first one corresponds to, "… tests described here rely on the relation of the empirical distribution function of the observations to the hypothesized distribution function in some manner". In this case, the distribution function is the theoretically adjusted one, id est the normal distribution with parameters equal to the mean and standard deviation of the logarithmic returns for prices. The null distribution is then constructed under the proposition that empirical data resembles the normal distribution.

Meanwhile, the second goodness of fit test uses the same approach, but with another perspective. These tests are constructed to compare to datasets. It could be understood as a generalization of the normality tests where the null hypothesis corresponds to “the samples come from the same theoretical distribution”. The tests are constructed to compare different qualities of the distributions and may lack statistical robustness. Hence, the general approach is then to perform a double check, the first one being the rejection of the normality assumption. The second is the confirmation of the goodness of fit for other (more appropriate) distribution.

This procedure’s quality means that the proposal is an excellent candidate to develop further models. The ability to capture a bubble expansion and burst can be translated to extreme quantile occurrences with higher probability than the one estimated with a normal distribution. With these results, it is possible to rethink some models relying on simulation procedures like all the Monte Carlo approaches to value derivative instruments or measure with greater precision the exposure and risk of financial institutions.

Once more, the problem of such models is the computational and parameter stability for complex distributions. The Expectation-Maximization algorithms are numerical methods that seek maxima in likelihood functions; the problem for such is that the increase in parameters takes a tradeoff: increases flexibility, but also increases the complexity of the function leading to cases where finding the optimal parameters becomes problematic, as no critical value may be found or by finding local instead of the global maximum. It is then necessary to test multiple starting points and test the convergence parameters to be stable. Furthermore, as these methods depend on the simulation algorithm, computational considerations must also be considered, mainly the pseudo-random number generation process. Any simulated vector would depend on it.

Another aspect to consider is the parameter stability through the series. As mentioned earlier in Chapter II’s remarks, where the methodology highly depends on the window length to perform the tests, a similar issue may occur in this case. At first glance, it is needed a large sample to converge to the population parameters, but this number is ambiguous and could be subject to various criteria. Furthermore, for financial series where stylized facts, such as volatility clusters, are typical, then the samples used to perform the parameter estimation algorithm can witness certain instability. However, this issue is part of almost any statistical study and maybe palliated with a robust test as the ones presented by Braun (1980). In practice, this procedure may extend to the actualization of datasets that maintain a certain constant length on the number of observations used. The result would be a rolling window with higher chances to be significant in the population parameter estimation.

6.3 CHAPTER IV

The chapter develops on using one of the statistical learning techniques to classify the bubbles into types or families. One of the most recent computational science aspects is the wide usage of numerical algorithms that can compute optimal parameters of problems with no close solution. It means that there is not a formula one can use for any problem. In practice, this means the usage of algorithms that deploy iterations into a convergence problem. Besides this modern capability to use such techniques with relative ease, implement some models born in the mid-XX century but practically available in XXI by the computational and software development (James, Witten, Hastie, Tibshirani, 2013, cap 2). One of the essential branches of these techniques is the classification models. Such problems were initially presented as a particularity of the traditional regression models, but with the difference that the response variable is a qualitative rather than quantitative measure. The model seeks to explain the chance of belonging to certain categories based on specific properties, mostly quantitative.

Many models deploy the classification algorithms from different perspectives, but a general categorization would identify the supervised and unsupervised models. The first ones refer to models that need thorough training based on previous classifications. Consequently, there is needed a dataset large enough to learn or detect patterns that explain the response. The closer type of these models with traditional regression is the logistic ones (James et al., 2013, cap 4). The traditional models refer to two classifications in the response variable, but it is possible to extend them to multiple regression models. However, some problems that may arise are related to the quality of the data. Because large data sets are needed, then a large portion of the population is required so that it is possible to make an inference; furthermore, as the model is a generalized case of the linear model, so if the relationship is no-linear, then the model will have a poor performance even when the variables do have a relationship.

Moreover, if the data is not balanced (the proportions are not equal), the algorithm tends to misfit certain variables' relevance. Finally, the general idea of the classification algorithms is to extrapolate the results and predict further observations. However, as there is no constraint in this property, it is usual to overfit the model to the data, losing any possibility of implementing the model in out-sample datasets. This property is also related to the instability of the parameters. When the classification variable is widely separated, the parameters once more loose statistical significance and are sensitive to small changes in the dataset.

Further developments use segmentation of the hyperspace so that classifications reside inside the boxes. The models that use such procedures are tree-based (James et al., 2013, cap 8). The mechanism of such depends on the segmentation of data into decision trees where conditional statements create planes that segment the hyperplane. These lines are translated into boxes that seek to create clusters that predict the observation's belonging to specific groups. These models are better at capturing no-linear structure in data. It does not depend on the specification provided, but instead uses an algorithm that chooses the space's best segmentation. Further advantages of the models include the ease to be explained as the interpretation of the decisions is direct; also, as they resemble human decision-making, it could be a natural way to make classification problems. The main issue in this type of model is that compared to other models, they usually underperform, so more complex models are needed.

A more flexible way to stratify spaces is proposed under the Support Vector Machines (SVM) (James et al., 2013, cap 9). These models are characterized to have one of the better performances in classification problems. The idea behind these is that observations can be divided by hyperplanes in the space. It implies the reduction of the dimension of the hyperspace in comparison to the available data. This line is intended to mark a frontier in which observations are efficiently divided. In most supervised learning problems, these models tend to be the better option because of the flexibility; nevertheless, there are some critical considerations to examine. The first and most important is the computational cost for these algorithms. Because the number of parameters depends on the dataset's dimension, and the convergence method is robust, the time elapsed for the numerical method to converge is, most of the time, considerable. Furthermore, the interpretability of the model is lost.

The final consideration of classification problems treated in this document is the unsupervised ones. The idea is that the datasets do not have a response variable, but only properties translated in numerical or categorical variables (James et al., 2013, cap 10). The idea is then to find the similarities and groups the information is center around, like the type of problem presented in Chapter IV. Such a study uses statistical criteria to determine the number of clusters that better describe the explanatory and response variable's distribution. For such, the algorithm then determines the groups and classifies the bubbles according to their similarity.

ANNEX

Table 24. S&P 500 index daily adjusted prices sample correlations.

AMZN	AMGN	ADBE	CAN	ABT	AABV	AAPL	
0.892	0.934	0.943	0.923	0.910	0.732	1.000	AAPL
0.780	0.739	0.761	0.749	0.680	1.000	0.732	AABV
0.934	0.905	0.966	0.952	1.000	0.680	0.910	ABT
0.942	0.896	0.969	1.000	0.952	0.749	0.923	CAN
0.979	0.906	1.000	0.969	0.966	0.761	0.943	ADBE
0.868	1.000	0.906	0.896	0.905	0.739	0.934	AMGN
1.000	0.868	0.979	0.942	0.934	0.780	0.892	AMZN
0.858	0.831	0.887	0.917	0.867	0.735	0.852	AVGO
0.853	0.796	0.857	0.883	0.804	0.859	0.832	BAC
- 0.198	0.056	- 0.149	- 0.117	- 0.162	0.100	0.014	BMY
0.908	0.848	0.907	0.928	0.866	0.860	0.858	BRK.B
0.736	0.759	0.777	0.853	0.797	0.549	0.788	CMCSA
0.909	0.901	0.949	0.940	0.958	0.603	0.935	COST
0.967	0.887	0.980	0.952	0.959	0.740	0.905	CRM
0.939	0.777	0.923	0.928	0.920	0.707	0.794	CSCO
0.747	0.620	0.717	0.795	0.669	0.717	0.610	CVX
0.674	0.701	0.738	0.809	0.805	0.387	0.739	DIS
0.846	0.790	0.860	0.896	0.789	0.806	0.834	FB
0.942	0.908	0.958	0.959	0.907	0.820	0.937	GOOG
0.939	0.907	0.955	0.955	0.902	0.821	0.936	GOOGL
0.940	0.893	0.959	0.978	0.935	0.770	0.926	HD
0.127	0.119	0.114	0.226	0.061	0.230	0.160	IBM
0.933	0.888	0.955	0.934	0.909	0.805	0.916	INTC
0.859	0.837	0.863	0.873	0.818	0.778	0.831	JNJ
0.896	0.864	0.915	0.943	0.869	0.841	0.898	JPM
0.809	0.853	0.876	0.921	0.897	0.568	0.864	KO
0.848	0.873	0.891	0.830	0.924	0.534	0.868	LLY
0.940	0.902	0.980	0.979	0.980	0.687	0.939	MA
0.903	0.839	0.932	0.957	0.931	0.699	0.859	MCD
0.820	0.894	0.868	0.905	0.884	0.552	0.878	MDT
0.835	0.863	0.877	0.885	0.920	0.486	0.839	MRK
0.940	0.920	0.980	0.961	0.968	0.684	0.958	MSFT
0.905	0.915	0.958	0.953	0.951	0.663	0.944	NEE
0.968	0.807	0.945	0.883	0.890	0.799	0.839	NFLX
0.880	0.865	0.918	0.907	0.937	0.593	0.879	NKE
0.862	0.786	0.841	0.813	0.729	0.910	0.838	NVDA
0.881	0.837	0.898	0.926	0.905	0.726	0.850	ORCL
0.828	0.857	0.881	0.919	0.897	0.567	0.885	PEP
0.863	0.759	0.831	0.817	0.842	0.691	0.706	PFE
0.733	0.818	0.811	0.848	0.864	0.411	0.826	PG
- 0.061	0.033	- 0.048	0.032	- 0.114	0.166	0.042	PM
0.962	0.868	0.979	0.971	0.972	0.756	0.898	PYPL
0.203	0.395	0.268	0.398	0.241	0.168	0.389	T
0.940	0.925	0.977	0.970	0.980	0.696	0.939	TMO
0.953	0.892	0.953	0.929	0.901	0.854	0.896	UNH
0.949	0.870	0.962	0.967	0.954	0.741	0.890	UNP
0.948	0.908	0.983	0.984	0.979	0.701	0.940	V
0.849	0.846	0.872	0.892	0.899	0.559	0.810	VZ
0.911	0.900	0.943	0.949	0.955	0.718	0.916	WMT
- 0.168	- 0.278	- 0.246	- 0.108	- 0.275	- 0.018	- 0.337	XOM

DIS	CVX	CSCO	CRM	COST	CMCSA	BRK-B	BMY	BAC	AVGO
0.739	0.610	0.794	0.905	0.935	0.788	0.858	0.014	0.832	0.852
0.387	0.717	0.707	0.740	0.603	0.549	0.860	0.100	0.859	0.735
0.805	0.669	0.920	0.959	0.958	0.797	0.866	- 0.162	0.804	0.867
0.809	0.795	0.928	0.952	0.940	0.853	0.928	- 0.117	0.883	0.917
0.738	0.717	0.923	0.980	0.949	0.777	0.907	- 0.149	0.857	0.887
0.701	0.620	0.777	0.887	0.901	0.759	0.848	0.056	0.796	0.831
0.674	0.747	0.939	0.967	0.909	0.736	0.908	- 0.198	0.853	0.858
0.726	0.823	0.872	0.865	0.827	0.906	0.919	- 0.220	0.896	1.000
0.694	0.867	0.854	0.845	0.750	0.796	0.963	- 0.149	1.000	0.896
- 0.151	- 0.259	- 0.326	- 0.139	- 0.112	- 0.208	- 0.146	1.000	- 0.149	- 0.220
0.699	0.879	0.909	0.899	0.813	0.827	1.000	- 0.146	0.963	0.919
0.807	0.735	0.777	0.752	0.806	1.000	0.827	- 0.208	0.796	0.906
0.805	0.603	0.854	0.925	1.000	0.806	0.813	- 0.112	0.750	0.827
0.755	0.730	0.939	1.000	0.925	0.752	0.899	- 0.139	0.845	0.865
0.754	0.828	1.000	0.939	0.854	0.777	0.909	- 0.326	0.854	0.872
0.601	1.000	0.828	0.730	0.603	0.735	0.879	- 0.259	0.867	0.823
1.000	0.601	0.754	0.755	0.805	0.807	0.699	- 0.151	0.694	0.726
0.636	0.829	0.811	0.821	0.766	0.820	0.888	- 0.105	0.896	0.920
0.708	0.781	0.881	0.941	0.898	0.823	0.937	- 0.079	0.915	0.924
0.706	0.781	0.877	0.938	0.893	0.822	0.936	- 0.081	0.915	0.922
0.771	0.785	0.903	0.943	0.936	0.845	0.925	- 0.100	0.890	0.912
0.216	0.481	0.207	0.097	0.076	0.400	0.321	- 0.020	0.396	0.328
0.665	0.736	0.879	0.936	0.896	0.730	0.893	- 0.087	0.869	0.880
0.590	0.785	0.824	0.830	0.786	0.841	0.913	- 0.084	0.834	0.914
0.755	0.852	0.882	0.896	0.841	0.844	0.977	- 0.117	0.981	0.923
0.829	0.657	0.800	0.864	0.915	0.863	0.811	- 0.039	0.729	0.849
0.707	0.469	0.793	0.893	0.912	0.698	0.729	- 0.109	0.641	0.752
0.825	0.700	0.915	0.969	0.972	0.820	0.880	- 0.145	0.833	0.883
0.777	0.791	0.913	0.914	0.899	0.883	0.910	- 0.213	0.847	0.940
0.811	0.641	0.789	0.862	0.916	0.823	0.785	- 0.026	0.713	0.813
0.820	0.649	0.849	0.874	0.918	0.855	0.791	- 0.267	0.708	0.844
0.775	0.647	0.876	0.948	0.982	0.813	0.860	- 0.107	0.804	0.876
0.765	0.652	0.852	0.927	0.970	0.845	0.853	- 0.093	0.786	0.891
0.605	0.722	0.915	0.945	0.831	0.632	0.875	- 0.215	0.844	0.809
0.820	0.597	0.860	0.941	0.935	0.703	0.787	- 0.064	0.738	0.768
0.430	0.737	0.739	0.791	0.709	0.649	0.883	- 0.016	0.885	0.810
0.760	0.789	0.897	0.884	0.850	0.900	0.913	- 0.182	0.869	0.949
0.812	0.654	0.810	0.847	0.926	0.931	0.815	- 0.114	0.751	0.896
0.662	0.768	0.911	0.880	0.744	0.660	0.861	- 0.214	0.768	0.758
0.803	0.535	0.730	0.777	0.905	0.880	0.701	- 0.134	0.625	0.821
- 0.065	0.237	- 0.045	- 0.103	- 0.110	0.348	0.169	0.118	0.191	0.288
0.763	0.757	0.950	0.968	0.926	0.815	0.915	- 0.196	0.861	0.910
0.384	0.374	0.199	0.210	0.379	0.606	0.346	0.175	0.329	0.450
0.787	0.687	0.904	0.958	0.972	0.835	0.881	- 0.114	0.815	0.897
0.652	0.806	0.899	0.936	0.857	0.766	0.960	- 0.137	0.913	0.897
0.806	0.814	0.965	0.968	0.913	0.834	0.941	- 0.232	0.899	0.914
0.817	0.720	0.922	0.970	0.971	0.835	0.895	- 0.151	0.842	0.897
0.746	0.675	0.855	0.877	0.914	0.787	0.806	- 0.136	0.685	0.798
0.745	0.675	0.871	0.908	0.951	0.845	0.878	- 0.105	0.805	0.886
- 0.063	0.426	0.017	- 0.163	- 0.341	- 0.031	0.041	- 0.086	0.046	- 0.043

LLY	KO	JPM	JNJ	INTC	IBM	HD	GOOGL	GOOG	FB
0.868	0.864	0.898	0.831	0.916	0.160	0.926	0.936	0.937	0.834
0.534	0.568	0.841	0.778	0.805	0.230	0.770	0.821	0.820	0.806
0.924	0.897	0.869	0.818	0.909	0.061	0.935	0.902	0.907	0.789
0.830	0.921	0.943	0.873	0.934	0.226	0.978	0.955	0.959	0.896
0.891	0.876	0.915	0.863	0.955	0.114	0.959	0.955	0.958	0.860
0.873	0.853	0.864	0.837	0.888	0.119	0.893	0.907	0.908	0.790
0.848	0.809	0.896	0.859	0.933	0.127	0.940	0.939	0.942	0.846
0.752	0.849	0.923	0.914	0.880	0.328	0.912	0.922	0.924	0.920
0.641	0.729	0.981	0.834	0.869	0.396	0.890	0.915	0.915	0.896
- 0.109	- 0.039	- 0.117	- 0.084	- 0.087	- 0.020	- 0.100	- 0.081	- 0.079	- 0.105
0.729	0.811	0.977	0.913	0.893	0.321	0.925	0.936	0.937	0.888
0.698	0.863	0.844	0.841	0.730	0.400	0.845	0.822	0.823	0.820
0.912	0.915	0.841	0.786	0.896	0.076	0.936	0.893	0.898	0.766
0.893	0.864	0.896	0.830	0.936	0.097	0.943	0.938	0.941	0.821
0.793	0.800	0.882	0.824	0.879	0.207	0.903	0.877	0.881	0.811
0.469	0.657	0.852	0.785	0.736	0.481	0.785	0.781	0.781	0.829
0.707	0.829	0.755	0.590	0.665	0.216	0.771	0.706	0.708	0.636
0.625	0.780	0.913	0.872	0.859	0.370	0.902	0.925	0.925	1.000
0.812	0.868	0.953	0.899	0.952	0.229	0.967	1.000	1.000	0.925
0.808	0.864	0.953	0.897	0.949	0.234	0.964	1.000	1.000	0.925
0.809	0.904	0.942	0.861	0.931	0.235	1.000	0.964	0.967	0.902
- 0.060	0.123	0.341	0.288	0.184	1.000	0.235	0.234	0.229	0.370
0.827	0.819	0.913	0.838	1.000	0.184	0.931	0.949	0.952	0.859
0.760	0.813	0.873	1.000	0.838	0.288	0.861	0.897	0.899	0.872
0.724	0.829	1.000	0.873	0.913	0.341	0.942	0.953	0.953	0.913
0.824	1.000	0.829	0.813	0.819	0.123	0.904	0.864	0.868	0.780
1.000	0.824	0.724	0.760	0.827	- 0.060	0.809	0.808	0.812	0.625
0.896	0.924	0.904	0.819	0.930	0.121	0.962	0.933	0.937	0.836
0.814	0.931	0.903	0.894	0.877	0.152	0.948	0.920	0.924	0.876
0.829	0.937	0.815	0.792	0.822	0.180	0.889	0.858	0.860	0.757
0.906	0.898	0.795	0.798	0.806	0.135	0.851	0.822	0.825	0.707
0.920	0.915	0.884	0.845	0.936	0.104	0.950	0.934	0.939	0.826
0.903	0.945	0.875	0.859	0.916	0.127	0.944	0.923	0.927	0.831
0.800	0.710	0.860	0.790	0.908	0.087	0.883	0.896	0.899	0.814
0.891	0.872	0.814	0.700	0.878	0.010	0.898	0.864	0.868	0.709
0.608	0.632	0.882	0.845	0.854	0.290	0.843	0.895	0.895	0.885
0.773	0.862	0.904	0.902	0.849	0.269	0.918	0.905	0.908	0.900
0.840	0.949	0.838	0.863	0.827	0.230	0.907	0.878	0.881	0.825
0.780	0.720	0.794	0.804	0.771	0.099	0.788	0.782	0.783	0.674
0.850	0.922	0.732	0.758	0.756	0.153	0.821	0.777	0.782	0.701
- 0.162	0.073	0.153	0.353	- 0.018	0.490	0.044	0.130	0.126	0.316
0.867	0.884	0.910	0.868	0.926	0.120	0.958	0.939	0.943	0.872
0.189	0.497	0.403	0.426	0.314	0.601	0.401	0.372	0.375	0.441
0.919	0.925	0.891	0.867	0.930	0.113	0.952	0.933	0.938	0.834
0.813	0.815	0.942	0.915	0.927	0.178	0.926	0.949	0.951	0.876
0.854	0.865	0.939	0.857	0.926	0.227	0.951	0.934	0.937	0.848
0.894	0.930	0.913	0.845	0.933	0.128	0.970	0.944	0.948	0.852
0.871	0.896	0.785	0.806	0.835	0.094	0.865	0.823	0.829	0.693
0.872	0.902	0.878	0.870	0.897	0.129	0.944	0.907	0.913	0.827
- 0.416	- 0.173	- 0.022	- 0.016	- 0.190	0.428	- 0.131	- 0.121	- 0.128	0.011

ORCL	NVDA	NKE	NFLX	NEE	MSFT	MRK	MDT	MCD	MA
0.850	0.838	0.879	0.839	0.944	0.958	0.839	0.878	0.859	0.939
0.726	0.910	0.593	0.799	0.663	0.684	0.486	0.552	0.699	0.687
0.905	0.729	0.937	0.890	0.951	0.968	0.920	0.884	0.931	0.980
0.926	0.813	0.907	0.883	0.953	0.961	0.885	0.905	0.957	0.979
0.898	0.841	0.918	0.945	0.958	0.980	0.877	0.868	0.932	0.980
0.837	0.786	0.865	0.807	0.915	0.920	0.863	0.894	0.839	0.902
0.881	0.862	0.880	0.968	0.905	0.940	0.835	0.820	0.903	0.940
0.949	0.810	0.768	0.809	0.891	0.876	0.844	0.813	0.940	0.883
0.869	0.885	0.738	0.844	0.786	0.804	0.708	0.713	0.847	0.833
- 0.182	- 0.016	- 0.064	- 0.215	- 0.093	- 0.107	- 0.267	- 0.026	- 0.213	- 0.145
0.913	0.883	0.787	0.875	0.853	0.860	0.791	0.785	0.910	0.880
0.900	0.649	0.703	0.632	0.845	0.813	0.855	0.823	0.883	0.820
0.850	0.709	0.935	0.831	0.970	0.982	0.918	0.916	0.899	0.972
0.884	0.791	0.941	0.945	0.927	0.948	0.874	0.862	0.914	0.969
0.897	0.739	0.860	0.915	0.852	0.876	0.849	0.789	0.913	0.915
0.789	0.737	0.597	0.722	0.652	0.647	0.649	0.641	0.791	0.700
0.760	0.430	0.820	0.605	0.765	0.775	0.820	0.811	0.777	0.825
0.900	0.885	0.709	0.814	0.831	0.826	0.707	0.757	0.876	0.836
0.908	0.895	0.868	0.899	0.927	0.939	0.825	0.860	0.924	0.937
0.905	0.895	0.864	0.896	0.923	0.934	0.822	0.858	0.920	0.933
0.918	0.843	0.898	0.883	0.944	0.950	0.851	0.889	0.948	0.962
0.269	0.290	0.010	0.087	0.127	0.104	0.135	0.180	0.152	0.121
0.849	0.854	0.878	0.908	0.916	0.936	0.806	0.822	0.877	0.930
0.902	0.845	0.700	0.790	0.859	0.845	0.798	0.792	0.894	0.819
0.904	0.882	0.814	0.860	0.875	0.884	0.795	0.815	0.903	0.904
0.862	0.632	0.872	0.710	0.945	0.915	0.898	0.937	0.931	0.924
0.773	0.608	0.891	0.800	0.903	0.920	0.906	0.829	0.814	0.896
0.900	0.761	0.948	0.891	0.973	0.984	0.916	0.910	0.941	1.000
0.946	0.764	0.849	0.840	0.937	0.925	0.882	0.863	1.000	0.941
0.825	0.639	0.882	0.709	0.924	0.904	0.914	1.000	0.863	0.910
0.853	0.587	0.864	0.752	0.923	0.912	1.000	0.914	0.882	0.916
0.886	0.782	0.926	0.879	0.985	1.000	0.912	0.904	0.925	0.984
0.891	0.757	0.902	0.832	1.000	0.985	0.923	0.924	0.937	0.973
0.827	0.856	0.842	1.000	0.832	0.879	0.752	0.709	0.840	0.891
0.801	0.641	1.000	0.842	0.902	0.926	0.864	0.882	0.849	0.948
0.788	1.000	0.641	0.856	0.757	0.782	0.587	0.639	0.764	0.761
1.000	0.788	0.801	0.827	0.891	0.886	0.853	0.825	0.946	0.900
0.904	0.677	0.835	0.723	0.955	0.931	0.918	0.915	0.924	0.918
0.805	0.666	0.783	0.843	0.746	0.768	0.794	0.730	0.812	0.805
0.822	0.525	0.808	0.617	0.924	0.893	0.931	0.890	0.856	0.875
0.230	0.227	- 0.231	- 0.111	0.029	- 0.046	- 0.021	0.028	0.118	- 0.082
0.934	0.806	0.902	0.929	0.942	0.957	0.876	0.844	0.962	0.974
0.371	0.275	0.237	0.058	0.439	0.361	0.412	0.517	0.359	0.329
0.915	0.769	0.926	0.879	0.981	0.988	0.931	0.915	0.942	0.985
0.890	0.902	0.826	0.930	0.892	0.907	0.817	0.808	0.907	0.907
0.917	0.790	0.906	0.915	0.920	0.936	0.898	0.862	0.936	0.963
0.914	0.777	0.940	0.894	0.975	0.985	0.919	0.915	0.954	0.998
0.814	0.607	0.880	0.758	0.912	0.901	0.923	0.884	0.875	0.898
0.912	0.776	0.872	0.836	0.961	0.961	0.887	0.879	0.934	0.950
- 0.064	- 0.094	- 0.239	- 0.164	- 0.311	- 0.333	- 0.193	- 0.140	- 0.105	- 0.245

V	UNP	UNH	TMO	T	PYPL	PM	PG	PFE	PEP
0.940	0.890	0.896	0.939	0.389	0.898	0.042	0.826	0.706	0.885
0.701	0.741	0.854	0.696	0.168	0.756	0.166	0.411	0.691	0.567
0.979	0.954	0.901	0.980	0.241	0.972	- 0.114	0.864	0.842	0.897
0.984	0.967	0.929	0.970	0.398	0.971	0.032	0.848	0.817	0.919
0.983	0.962	0.953	0.977	0.268	0.979	- 0.048	0.811	0.831	0.881
0.908	0.870	0.892	0.925	0.395	0.868	0.033	0.818	0.759	0.857
0.948	0.949	0.953	0.940	0.203	0.962	- 0.061	0.733	0.863	0.828
0.897	0.914	0.897	0.897	0.450	0.910	0.288	0.821	0.758	0.896
0.842	0.899	0.913	0.815	0.329	0.861	0.191	0.625	0.768	0.751
- 0.151	- 0.232	- 0.137	- 0.114	0.175	- 0.196	0.118	- 0.134	- 0.214	- 0.114
0.895	0.941	0.960	0.881	0.346	0.915	0.169	0.701	0.861	0.815
0.835	0.834	0.766	0.835	0.606	0.815	0.348	0.880	0.660	0.931
0.971	0.913	0.857	0.972	0.379	0.926	- 0.110	0.905	0.744	0.926
0.970	0.968	0.936	0.958	0.210	0.968	- 0.103	0.777	0.880	0.847
0.922	0.965	0.899	0.904	0.199	0.950	- 0.045	0.730	0.911	0.810
0.720	0.814	0.806	0.687	0.374	0.757	0.237	0.535	0.768	0.654
0.817	0.806	0.652	0.787	0.384	0.763	- 0.065	0.803	0.662	0.812
0.852	0.848	0.876	0.834	0.441	0.872	0.316	0.701	0.674	0.825
0.948	0.937	0.951	0.938	0.375	0.943	0.126	0.782	0.783	0.881
0.944	0.934	0.949	0.933	0.372	0.939	0.130	0.777	0.782	0.878
0.970	0.951	0.926	0.952	0.401	0.958	0.044	0.821	0.788	0.907
0.128	0.227	0.178	0.113	0.601	0.120	0.490	0.153	0.099	0.230
0.933	0.926	0.927	0.930	0.314	0.926	- 0.018	0.756	0.771	0.827
0.845	0.857	0.915	0.867	0.426	0.868	0.353	0.758	0.804	0.863
0.913	0.939	0.942	0.891	0.403	0.910	0.153	0.732	0.794	0.838
0.930	0.865	0.815	0.925	0.497	0.884	0.073	0.922	0.720	0.949
0.894	0.854	0.813	0.919	0.189	0.867	- 0.162	0.850	0.780	0.840
0.998	0.963	0.907	0.985	0.329	0.974	- 0.082	0.875	0.805	0.918
0.954	0.936	0.907	0.942	0.359	0.962	0.118	0.856	0.812	0.924
0.915	0.862	0.808	0.915	0.517	0.844	0.028	0.890	0.730	0.915
0.919	0.898	0.817	0.931	0.412	0.876	- 0.021	0.931	0.794	0.918
0.985	0.936	0.907	0.988	0.361	0.957	- 0.046	0.893	0.768	0.931
0.975	0.920	0.892	0.981	0.439	0.942	0.029	0.924	0.746	0.955
0.894	0.915	0.930	0.879	0.058	0.929	- 0.111	0.617	0.843	0.723
0.940	0.906	0.826	0.926	0.237	0.902	- 0.231	0.808	0.783	0.835
0.777	0.790	0.902	0.769	0.275	0.806	0.227	0.525	0.666	0.677
0.914	0.917	0.890	0.915	0.371	0.934	0.230	0.822	0.805	0.904
0.926	0.873	0.816	0.936	0.547	0.893	0.183	0.959	0.696	1.000
0.819	0.886	0.872	0.816	0.074	0.854	- 0.072	0.610	1.000	0.696
0.876	0.805	0.714	0.897	0.563	0.821	0.085	1.000	0.610	0.959
- 0.046	- 0.026	0.077	- 0.016	0.511	- 0.025	1.000	0.085	- 0.072	0.183
0.979	0.969	0.937	0.970	0.243	1.000	- 0.025	0.821	0.854	0.893
0.339	0.296	0.261	0.360	1.000	0.243	0.511	0.563	0.074	0.547
0.988	0.952	0.915	1.000	0.360	0.970	- 0.016	0.897	0.816	0.936
0.919	0.943	1.000	0.915	0.261	0.937	0.077	0.714	0.872	0.816
0.966	1.000	0.943	0.952	0.296	0.969	- 0.026	0.805	0.886	0.873
1.000	0.966	0.919	0.988	0.339	0.979	- 0.046	0.876	0.819	0.926
0.904	0.887	0.828	0.918	0.430	0.877	- 0.082	0.883	0.813	0.882
0.956	0.923	0.901	0.969	0.379	0.951	0.014	0.887	0.777	0.929
- 0.224	- 0.070	- 0.098	- 0.263	0.094	- 0.177	0.278	- 0.310	0.104	- 0.223

XOM	WMT	VZ
- 0.337	0.916	0.810
- 0.018	0.718	0.559
- 0.275	0.955	0.899
- 0.108	0.949	0.892
- 0.246	0.943	0.872
- 0.278	0.900	0.846
- 0.168	0.911	0.849
- 0.043	0.886	0.798
0.046	0.805	0.685
- 0.086	- 0.105	- 0.136
0.041	0.878	0.806
- 0.031	0.845	0.787
- 0.341	0.951	0.914
- 0.163	0.908	0.877
0.017	0.871	0.855
0.426	0.675	0.675
- 0.063	0.745	0.746
0.011	0.827	0.693
- 0.128	0.913	0.829
- 0.121	0.907	0.823
- 0.131	0.944	0.865
0.428	0.129	0.094
- 0.190	0.897	0.835
- 0.016	0.870	0.806
- 0.022	0.878	0.785
- 0.173	0.902	0.896
- 0.416	0.872	0.871
- 0.245	0.950	0.898
- 0.105	0.934	0.875
- 0.140	0.879	0.884
- 0.193	0.887	0.923
- 0.333	0.961	0.901
- 0.311	0.961	0.912
- 0.164	0.836	0.758
- 0.239	0.872	0.880
- 0.094	0.776	0.607
- 0.064	0.912	0.814
- 0.223	0.929	0.882
0.104	0.777	0.813
- 0.310	0.887	0.883
0.278	0.014	- 0.082
- 0.177	0.951	0.877
0.094	0.379	0.430
- 0.263	0.969	0.918
- 0.098	0.901	0.828
- 0.070	0.923	0.887
- 0.224	0.956	0.904
- 0.119	0.906	1.000
- 0.274	1.000	0.906
1.000	- 0.274	- 0.119

ANNEX II

Table 25. Augmented Dickey-Fuller test for prices (p-values).

Period	AMZN	UNH	NVDA	DIS	ADBE	NEE	NKE	UNP
1P	0.724	0.9207	0.5338	0.5873	0.4148	0.5336	0.1856	0.42
2P	0.9258	0.495	0.8744	0.1071	0.8956	0.8485	0.06766	0.2065

Table 26. Augmented Dickey-Fuller test for returns (p-values).

Period	AMZN	UNH	NVDA	DIS	ADBE	NEE	NKE	UNP
1P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 27. Phillips-Perron test for prices (p-values).

Period	AMZN	UNH	NVDA	DIS	ADBE	NEE	NKE	UNP
1P	0.3714	0.9229	0.5629	0.5146	0.3714	0.3667	0.1757	0.2789
2P	0.8631	0.2494	0.7976	0.3776	0.8007	0.5477	0.03263	0.1669

Table 28. Phillips-Perron test for returns (p-values).

Period	AMZN	UNH	NVDA	DIS	ADBE	NEE	NKE	UNP
1P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 29. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for prices (p-value).

Period	AMZN	UNH	NVDA	DIS	ADBE	NEE	NKE	UNP
1P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2P	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 30. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for returns (p-value).

Period	AMZN	UNH	NVDA	DIS	ADBE	NEE	NKE	UNP
1P	0.09458	0.05068	0.1	0.01	0.01	0.1	0.1	0.1
2P	0.1	0.1	0.07832	0.01	0.01	0.1	0.1	0.1

ANNEX III

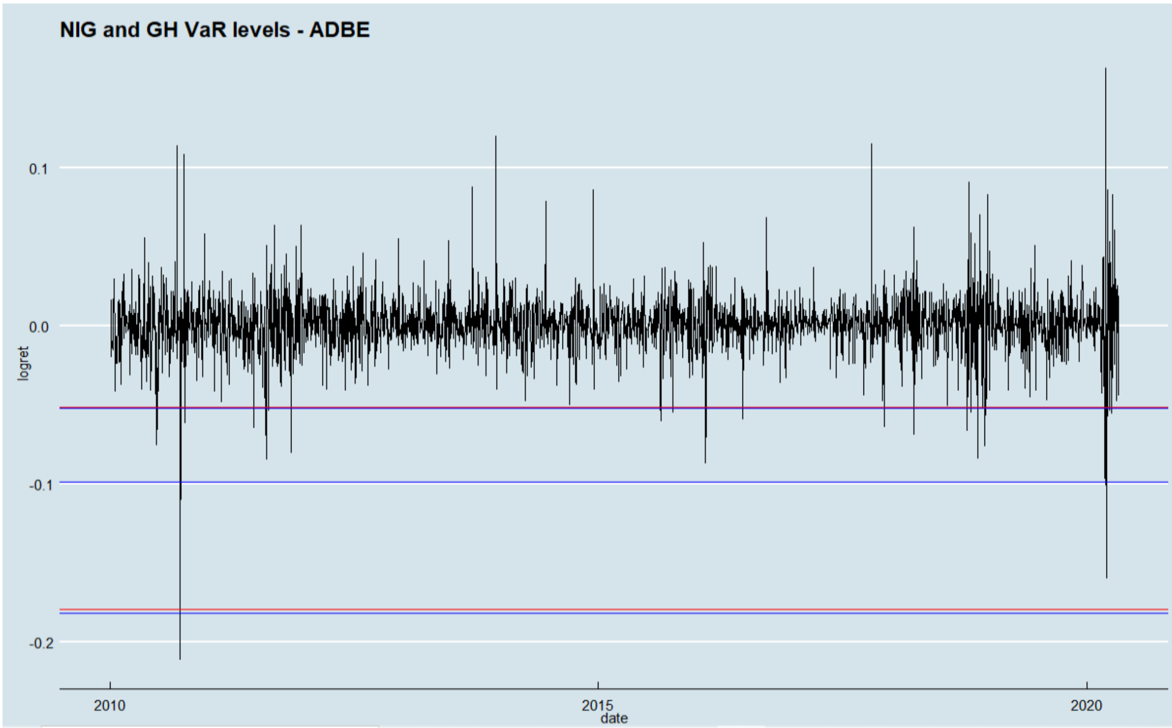

Fig. 40 NIG and GH VaR levels – ADBE.

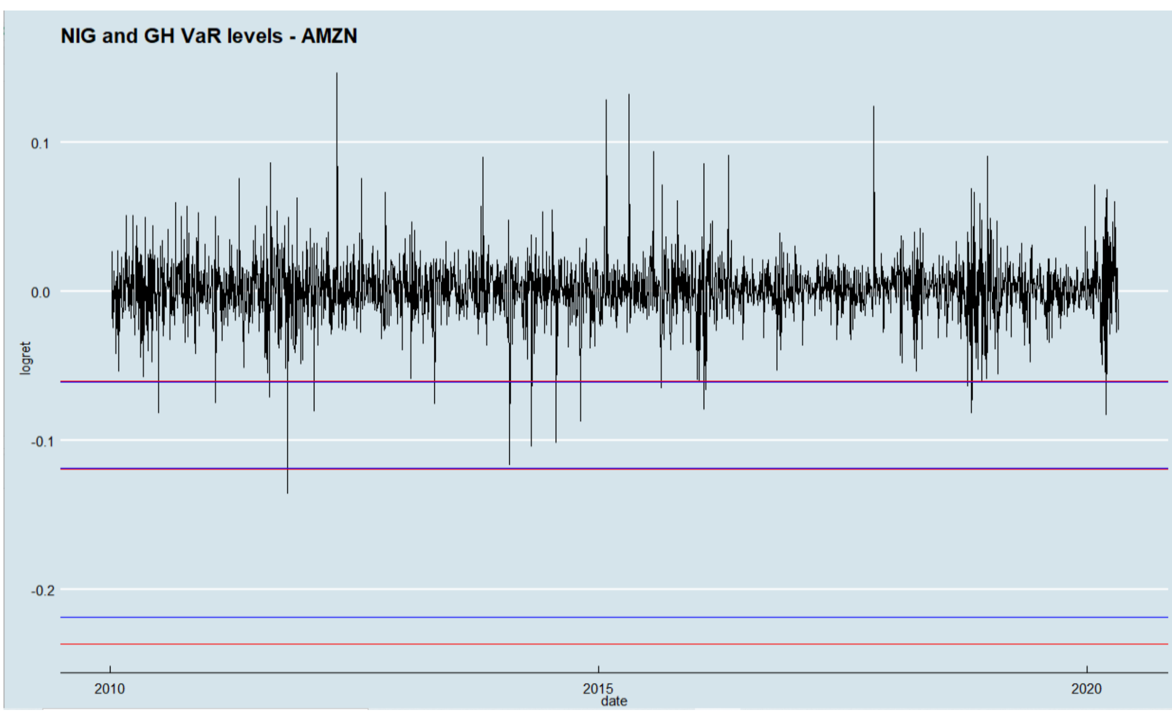

Fig. 41 NIG and GH VaR levels – AMZN.

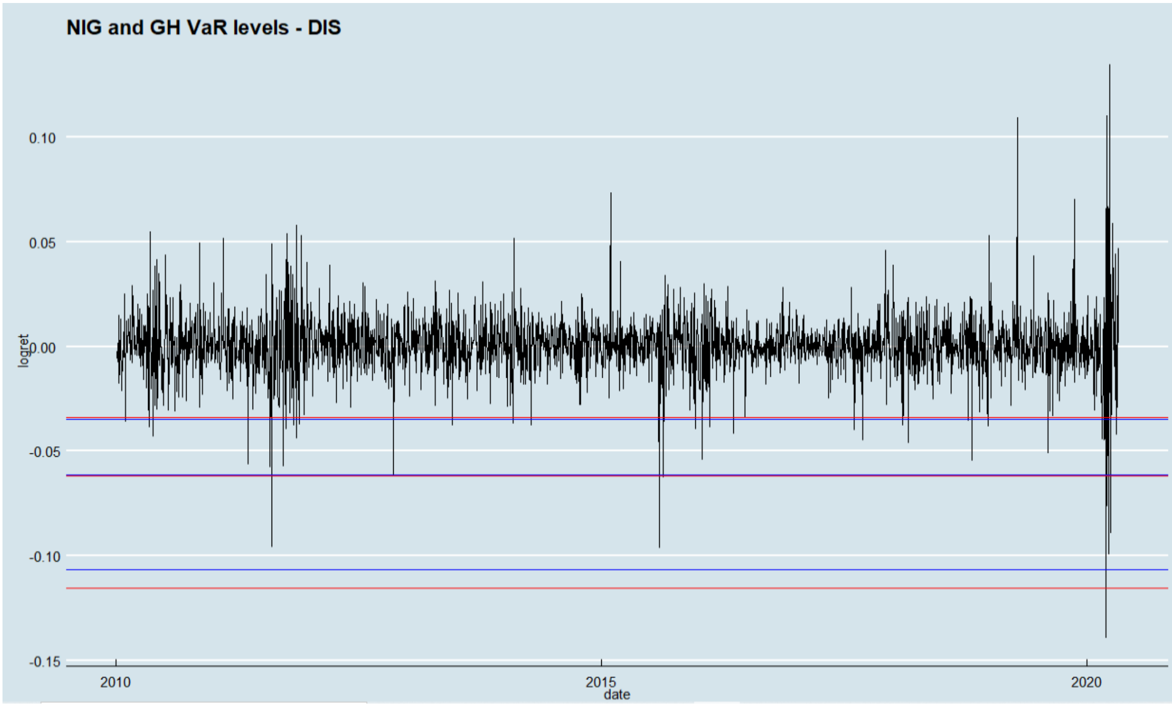

Fig. 42 NIG and GH VaR levels – DIS.

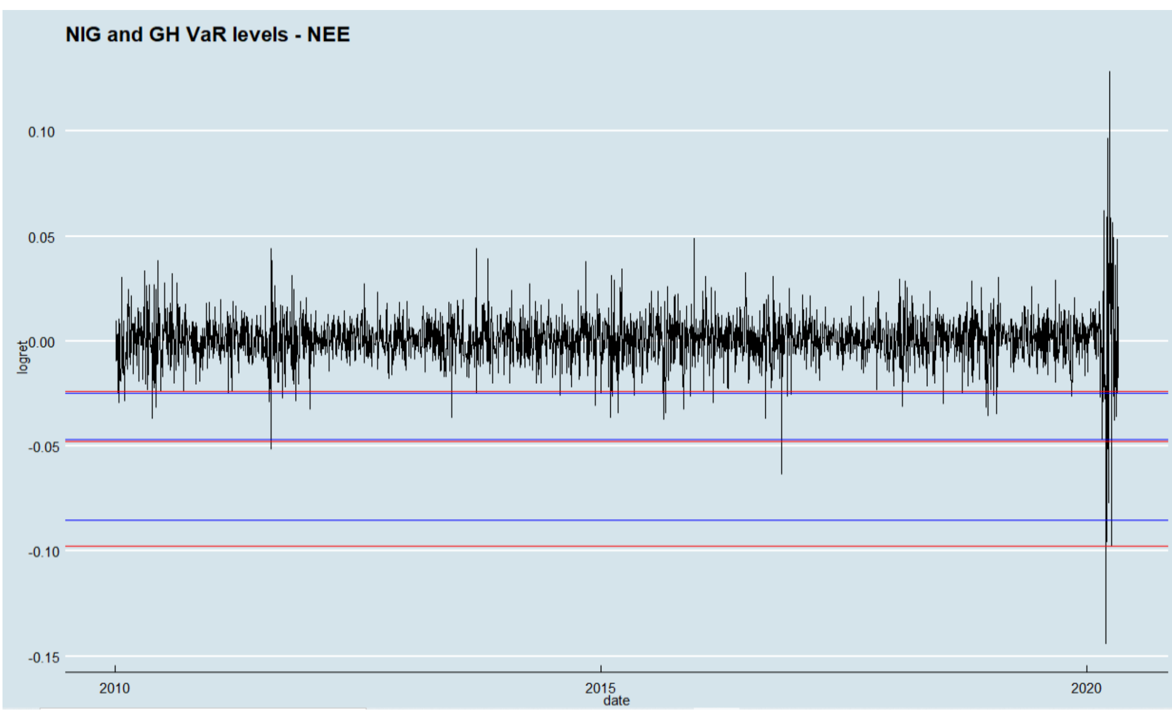

Fig. 43 NIG and GH VaR levels – NEE.

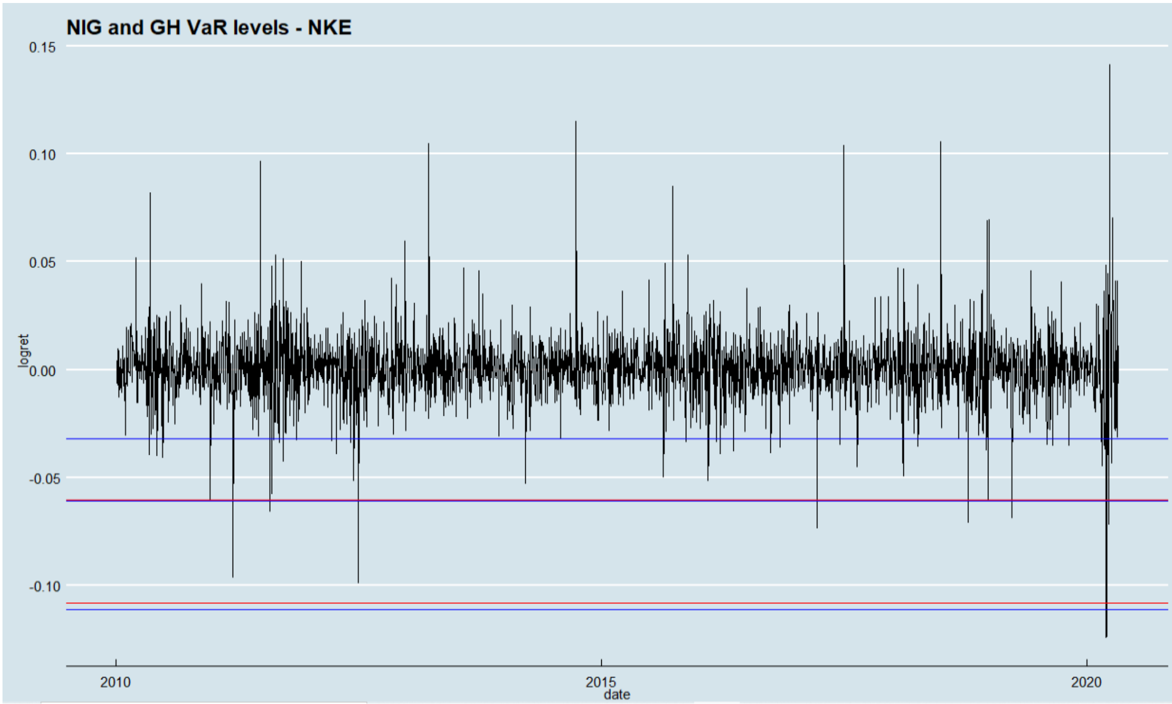

Fig. 44 NIG and GH VaR levels – NKE.

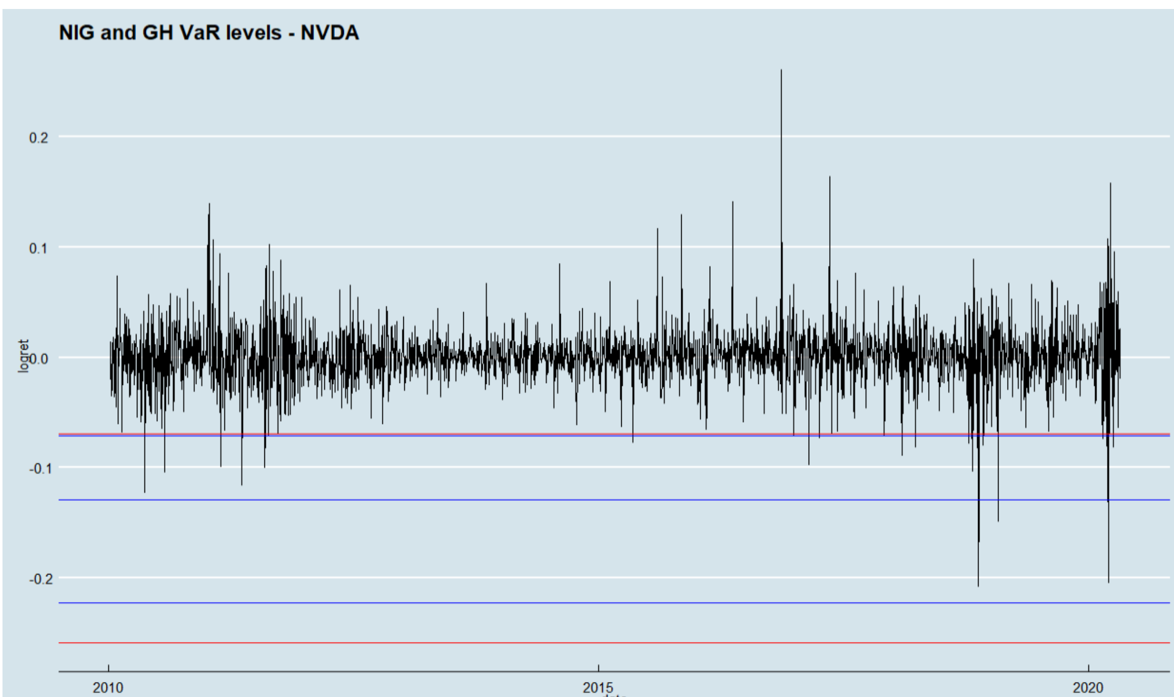

Fig. 45 NIG and GH VaR levels – NVDA.

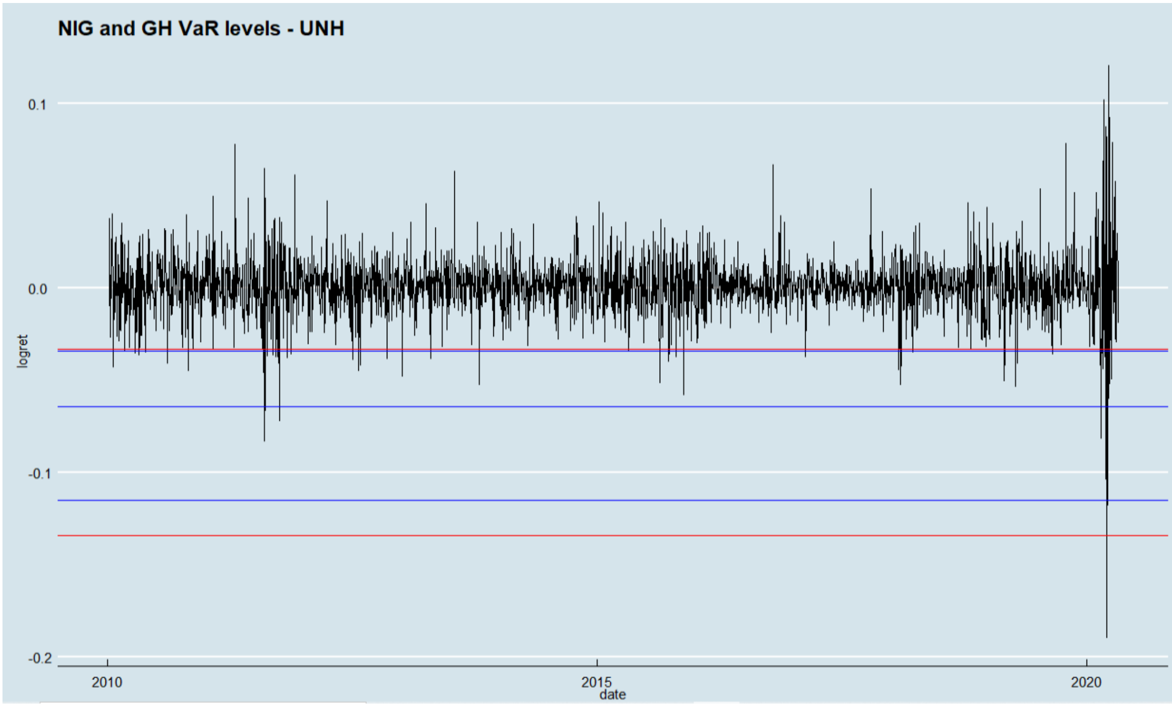

Fig. 46 NIG and GH VaR levels – UNH.

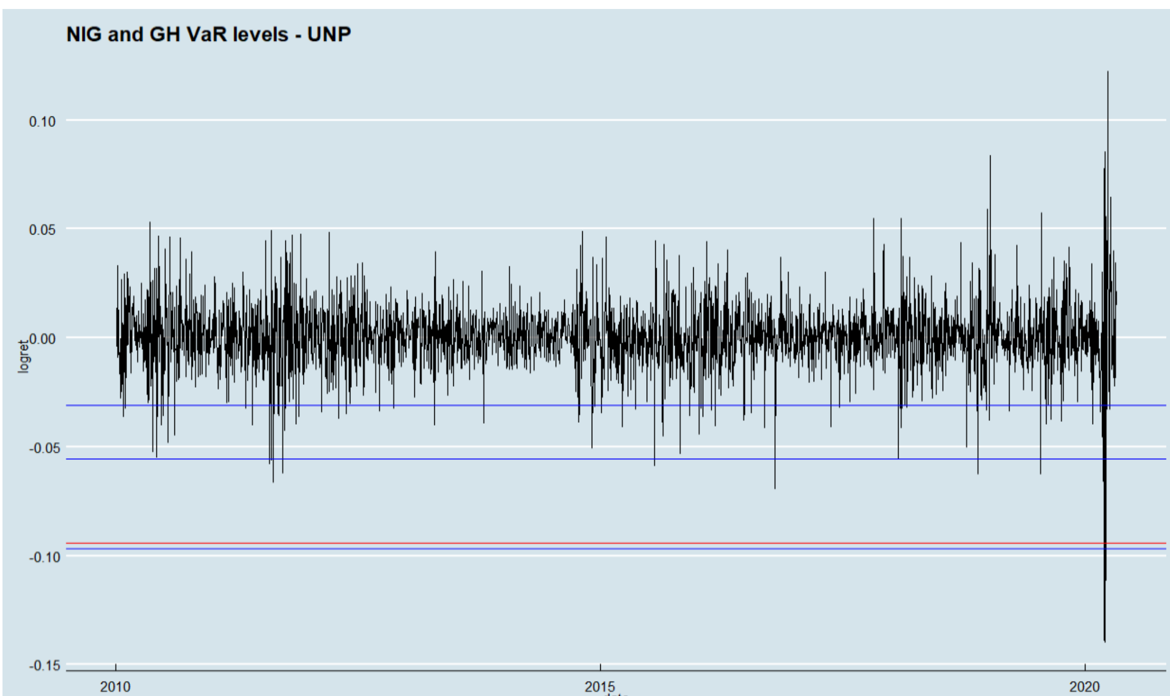

Fig. 47 NIG and GH VaR levels – UNP.

REFERENCES

Abreu, D., & Brunnermeier, M. K. (2003). Bubbles and crashes. *Econometrica, 71*(1), 173-204.

Albrecher, H., Ladoucette, S. A., & Schoutens, W. (2007). A generic one-factor Lévy model for pricing synthetic CDOs. In Advances in Mathematical Finance (pp. 259-277). Birkhäuser Boston.

Albrecher, H., & Predota, M. (2004). On Asian option pricing for NIG Lévy processes. Journal of computational and applied mathematics, 172(1), 153-168.

Aliber, R. Z., & Kindleberger, C. P. (2017). *Manias, panics, and crashes: a history of financial crises*. Springer.

Allen, F., & Gale, D. (2000). Bubbles and crises. *The economic journal*, *110*(460), 236-255.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, *353*(1674), 401-419.

Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*, 803-821.

Benth, F. E., Groth, M., & Kettler, P. C. (2006). A quasi-Monte Carlo algorithm for the normal inverse Gaussian distribution and valuation of financial derivatives. International Journal of Theoretical and Applied Finance, 9(06), 843-867.

Blanchard, O. J., & Watson, M. W. (1982). *Bubbles, rational expectations and financial markets* (No. w0945). National Bureau of economic research.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The quarterly journal of economics*, *98*(1), 85-106.

Beyerer, J., Richter, M., & Nagel, M. (2017). *Pattern recognition: introduction, features, classifiers and principles*. Walter de Gruyter GmbH & Co KG.

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of economic Perspectives*, *29*(2), 213-38.

Bose, I., & Pal, R. (2006). Predicting the survival or failure of click-and-mortar corporations: A knowledge discovery approach. *European Journal of Operational Research*, *174*(2), 959-982.

Boyarchenko, S., & Levendorskiĭ, S. (2002). Barrier options and touch-and-out options under regular Lévy processes of exponential type. The Annals of Applied Probability, 12(4), 1261-1298.

Boyer, R. (2000). Is a finance-led growth regime a viable alternative to Fordism? A preliminary analysis. *Economy and society*, *29*(1), 111-145.

Brown, J. S., & Duguid, P. (2000). Balancing act: How to capture knowledge without killing it. *Harvard business review*, *78*(3), 73-73.

Browne, R. P., & McNicholas, P. D. (2015). A mixture of generalized hyperbolic distributions. *Canadian Journal of Statistics*, *43*(2), 176-198.

Camerer, C. (1989). Bubbles and fads in asset prices. *Journal of Economic Surveys*, *3*(1), 3-41.

Campbell, J. Y., Champbell, J. J., Campbell, J. W., Lo, A. W., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. princeton University press.

Cao, J., Ruan, X., Su, S., & Zhang, W. (2020). Pricing VIX derivatives with infinite-activity jumps. Journal of Futures Markets, 40(3), 329-354.

Caspi, I. (2013). Rtadf: Testing for bubbles with EViews. *Journal of Statistical Software*.

Caspi, I., Katzke, N., & Gupta, R. (2018). Date stamping historical periods of oil price explosivity: 1876–2014. *Energy Economics*, *70*, 582-587.

Celeux, G., & Govaert, G. (1993). Gaussian parsimonious clustering models. *Pattern Recognition, 28*, 781-793.

Cerecedo-Hernández, D., Franco-Ruiz, C. A., Contreras-Valdez, M. I., & Franco-Ruiz, J. A. (2019). Explosion in Virtual Assets (Cryptocurrencies). *Revista Mexicana de Economía y Finanzas (REMEF): nueva época, 14*(4), 715-727.

Chang, T., Gil-Alana, L., Aye, G. C., Gupta, R., & Ranjbar, O. (2016). Testing for bubbles in the BRICS stock markets. *Journal of Economic Studies*.

Chang, Y. P., Hung, M. C., Liu, H., & Jan, J. F. (2005). Testing symmetry of a NIG distribution. Communications in Statistics—Simulation and Computation®, 34(4), 851-862.

Chen, J., Hong, H., & Stein, J. C. (2001). Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of financial Economics*, *61*(3), 345-381.

Chuang, S. L., & Brockett, P. L. (2014). Modeling and pricing longevity derivatives using stochastic mortality rates and the esscher transform. North American Actuarial Journal, 18(1), 22-37.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1, 223-226.

Day, C. C. (2004). Is there a tulip in your future?: ruminations on tulip mania and the innovative dutch futures markets. *Journal des Economistes et des Etudes Humaines*, *14*(2), 151-170.

De la Vega, J., & Kellenbenz, H. (1957). *Confusion de Confusiones. Portions Descriptive of the Amsterdam Stock Exchange Selected and Translated by Hermann Kellenbenz*. Baker Library, Harvard Graduate School of Business Administration.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *the Journal of Finance*, *45*(2), 379-395.

Deringer, W. (2015). For what it's worth: historical financial bubbles and the boundaries of economic rationality. *Isis*, *106*(3), 646-656.

Diba, B. T., & Grossman, H. I. (1988). Explosive rational bubbles in stock prices?. *The American Economic Review*, *78*(3), 520-530.

Durlauf, S. N., & Blume, L. (Eds.). (2008). *The new Palgrave dictionary of economics* (Vol. 6). London, UK: Palgrave Macmillan.

Eberlein, E., & Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, *1*(3), 281-299.

Eberlein, E., & Prause, K. (2002). The generalized hyperbolic model: financial derivatives and risk measures. In *Mathematical Finance—Bachelier Congress 2000* (pp. 245-267). Springer, Berlin, Heidelberg.

Efendi, J., Srivastava, A., & Swanson, E. P. (2007). Why do corporate managers misstate financial statements? The role of option compensation and other factors. *Journal of financial economics*, *85*(3), 667-708.

Engel, C. (1996). The forward discount anomaly and the risk premium: A survey of recent evidence. *Journal of empirical finance*, *3*(2), 123-192.

Escobari, D., Garcia, S., & Mellado, C. (2017). Identifying bubbles in Latin American equity markets: Phillips-Perron-based tests and linkages. *Emerging Markets Review*, *33*, 90–101.

Escobari, D., & Jafarinejad, M. (2016). Date stamping bubbles in real estate investment trusts. *The Quarterly Review of Economics and Finance*, *60*, 224-230.

Evans, G. W. (1986). A test for speculative bubbles in the sterling-dollar exchange rate: 1981-84. *The American Economic Review*, 621-636.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review, 81*(4), 922-930.

Flood, R. P., & Garber, P. M. (1980). Market fundamentals versus price-level bubbles: the first tests. Journal of political economy, 88(4), 745-770.

Forsberg, L., & Bollerslev, T. (2002). Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model. Journal of Applied Econometrics, 17(5), 535-548.

Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for model-based cluster analysis. *Journal of Classification, 16*(2), 297-306.

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, *97*(458), 611-631.

Frestad, D., Benth, F. E., & Koekebakker, S. (2010). Modeling term structure dynamics in the Nordic electricity swap market. The Energy Journal, 31(2).

Friedman, M., & Schwartz, A. J. (2008). *A monetary history of the United States, 1867-1960*. Princeton University Press.

Gan, G., Ma, C., & Wu, J. (2007). *Data clustering: theory, algorithms, and applications* (Vol. 20). Siam.

Gerber, H. U., & Shiu, E. S. (1993). Option pricing by Esscher transforms. HEC Ecole des hautes études commerciales.

Gerlach, J. C., Demos, G., & Sornette, D. (2019). Dissection of Bitcoin's multiscale bubble history from January 2012 to February 2018. *Royal Society open science*, *6*(7), 180643.

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

Godin, F., Mayoral, S., & Morales, M. (2012). Contingent claim pricing using a normal inverse Gaussian probability distortion operator. Journal of Risk and Insurance, 79(3), 841-866.

Gordon, A. D. (1981). *Classification: methods for the exploratory analysis of multivariate data*. Chapman & Hall, London.

Griffin, D. (2014). Bitcoin: No Trust Needed. *Information Today*, *31*(4), 1.

Hainaut, D., & MacGilchrist, R. (2010). An interest rate tree driven by a Lévy process. The Journal of Derivatives, 18(2), 33-45.

Hamilton, J. D. (1994). *Time series analysis* (Vol. 2). New Jersey: Princeton.

Hartigan, J. A. (1975). *Clustering algorithms*. John Wiley & Sons, Inc..

Harvey, D. I., Leybourne, S. J., Sollis, R., & Taylor, A. R. (2016). Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance*, *38*, 548-574.

Harvey, D. I., Leybourne, S. J., & Zu, Y. (2020). Sign-based unit root tests for explosive financial bubbles in the presence of deterministically time-varying volatility. *Econometric Theory*, *36*(1), 122-169.

Heilbroner, R., & Milberg, W. (2008). *La evolución de la sociedad económica.* Pearson Educación.

Hofer, M., & Mayer, P. (2013). Pricing and hedging of lookback options in hyper-exponential jump diffusion models. Applied Mathematical Finance, 20(5), 489-511.

Homm, U., & Breitung, J. (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics, 10*(1), 198-231.

Hu, Y., & Oxley, L. (2017). Are there bubbles in exchange rates? Some new evidence from G10 and emerging market economies. *Economic Modelling*, *64*, 419–442.

Hu, Y., & Oxley, L. (2018). Do 18th century 'bubbles' survive the scrutiny of 21st century time series econometrics?. *Economics Letters*, *162*, 131-134.

James Gareth, Witten Daniela, Hastie Trevor & Tibshirani Robert (3013). An Introduction to Statistical Learning. Springer Texts in Statistics

Jeannin, M., & Pistorius, M. (2010). A transform approach to compute prices and Greeks of barrier options driven by a class of Lévy processes. Quantitative Finance, 10(6), 629-644.

Jiang, Z. Q., Zhou, W. X., Sornette, D., Woodard, R., Bastiaensen, K., & Cauwels, P. (2010). Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. *Journal of economic behavior & organization*, *74*(3), 149-162.

Johansen, A., & Sornette, D. (1999). Critical crashes. *Risk, 12,* 91-94.

Kaizoji, T., Leiss, M., Saichev, A., & Sornette, D. (2015). Super-exponential endogenous bubbles in an equilibrium model of fundamentalist and chartist traders. *Journal of Economic Behavior & Organization*, *112*, 289-310.

Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation of the normal–inverse Gaussian distribution. *Statistics & probability letters*, *57*(1), 43-52.

Karlis, D., & Lillestöl, J. (2004). Bayesian estimation of NIG models via Markov chain Monte Carlo methods. Applied Stochastic Models in Business and Industry, 20(4), 323-338.

Kassberger, S., & Kiesel, R. (2006). A fully parametric approach to return modelling and risk management of hedge funds. Financial markets and portfolio management, 20(4), 472-491.

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344). John Wiley & Sons.

Kiliç, R. (2007). Conditional volatility and distribution of exchange rates: GARCH and FIGARCH models with NIG distribution. Studies in Nonlinear Dynamics & Econometrics, 11(3).

Kirkby, J. L. (2017). Robust barrier option pricing by frame projection under exponential Lévy dynamics. Applied Mathematical Finance, 24(4), 337-386.

Koedijk, K. G., Schafgans, M. M., & De Vries, C. G. (1990). The tail index of exchange rate returns. *Journal of international economics*, *29*(1-2), 93-108.

Kräussl, R., Lehnert, T., & Martelin, N. (2016). Is there a bubble in the art market? *Journal of Empirical Finance*, *35*, 99–109.

Kumar, S., & Kumar, S. (2008). Collaboration in research productivity in oil seed research institutes of India. In *Proceedings of Fourth International Conference on Webometrics, Informetrics and Scientometrics* (pp. 28-1).

Kurozumi, E. (2020). Asymptotic properties of bubble monitoring tests. *Econometric Reviews, 39*(5), 510-538.

Li, Z. Z., Tao, R., Su, C. W., & Lobonţ, O. R. (2019). Does Bitcoin bubble burst?. *Quality & Quantity*, *53*(1), 91-105.

Lee, J. H., & Phillips, P. C. B. (2016). Asset pricing with financial bubble risk. *Journal of Empirical Finance*, *38*, 590-622.

Long, W., Li, D., & Li, Q. (2016). Testing Explosive Behavior in the Gold Market. *Empirical Economics*, *51*(3), 1151–1164.

Lux, T. (1998). The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions. *Journal of Economic Behavior & Organization*, *33*(2), 143-165.

Lux, T., & Sornette, D. (2002). On rational bubbles and fat tails. *Journal of Money, Credit and Banking*, 589-610.

McLachlan, G. J., & Basford, K. E. (1988). *Mixture models: Inference and applications to clustering* (Vol. 38). New York: M. Dekker.

Meese, R. A. (1986). Testing for bubbles in exchange markets: A case of sparkling rates?. *Journal of political economy*, *94*(2), 345-373.

Mizen, P. (2008). The credit crunch of 2007-2008: a discussion of the background, market reactions, and policy responses. *Federal Reserve Bank of St. Louis Review, 90* (September/October 2008).

Mobasher, B. (2018). *Origins the story of everything*. California, United States of America: Cognella Academic Publishing.

Monschang, V., & Wilfling, B. (2020). Sup-ADF-style bubble-detection methods under test. *Empirical Economics: Journal of the Institute for Advanced Studies, Vienna, Austria*, 1.

Mossavar-Rahmani, S., Nelson, B., Weir, M., Minovi, M., Ubide, A., Asl, F., Dibo, M., and Rich, M. C. (2018). (Un) Steady as She Goes. Technical Report January, Goldman Sachs Investment Management Division.

Nakamoto, S. (2009). *Bitcoin: A peer-to-peer electronic cash system*.

Neal, L. (1993). *The rise of financial capitalism: International capital markets in the age of reason*. Cambridge University Press.

Núñez, J. A., Contreras-Valdez, M. I., & Franco-Ruiz, C. A. (2019). Statistical analysis of bitcoin during explosive behavior periods. *PloS one*, *14*(3).

Ofek, E., & Richardson, M. (2003). Dotcom mania: The rise and fall of internet stock prices. *The Journal of Finance*, *58*(3), 1113-1137.

Önalan, Ö. (2009). Martingale measures for NIG Lévy processes with applications to mathematical finance. International Research Journal of Finance and Economics, 34, 56-68.

Pagan, A. (1996). The econometrics of financial markets. *Journal of empirical finance*, *3*(1), 15-102.

Panko, R. R. (2008). IT employment prospects: beyond the dotcom bubble. *European Journal of Information Systems*, *17*(3), 182-197.

Pavlidis, E., Yusupova, A., Paya, I., Peel, D., Martínez-García, E., Mack, A., & Grossman, V. (2016). Episodes of exuberance in housing markets: in search of the smoking gun. *The Journal of Real Estate Finance and Economics*, *53*(4), 419-449.

Phillips, P. C. B. (2016). Modeling speculative bubbles with diverse investor expectations. *Research in Economics*, *70*(3), 375-387.

Phillips, P. C., & Shi, S. P. (2018). Financial bubble implosion and reverse regression. *Econometric Theory*, *34*(4), 705-753.

Phillips, P. C., & Shi, S. P. (2019). Detecting financial collapse and ballooning sovereign risk. *Oxford Bulletin of Economics and Statistics, 81*(6), 1336-1361.

Phillips, P. C. B., Shi, S., & Yu, J. (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics*, *76*(3), 315-333.

Phillips, P. C. B., Shi, S., & Yu, J. (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International economic review*, *56*(4), 1043-1078.

Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?. *International economic review*, *52*(1), 201-226.

Phillips, P. C. B., & Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics*, *2*(3), 455-491.

Rasmus, S., Asmussen, S., & Wiktorsson, M. (2004, June). Pricing of some exotic options with NIG-Lévy input. In International Conference on Computational Science (pp. 795-802). Springer, Berlin, Heidelberg.

Razi, M. A., Tarn, J. M., & Siddiqui, F. A. (2004). Exploring the failure and success of DotComs. *Information Management & Computer Security*.

Ribstein, L. E. (2003). Bubble laws. *Hous. L. Rev.*, *40*, 77.

Romer, C. D. (1990). The great crash and the onset of the great depression. *The Quarterly Journal of Economics*, *105*(3), 597-624.

Rydberg, T. H. (1999). Generalized hyperbolic diffusion processes with applications in finance. *Mathematical Finance*, *9*(2), 183-201.

Sarna, D. E. (2010). *History of greed: Financial fraud from tulip mania to Bernie Madoff*. John Wiley & Sons.

Schachter, S., Gerin, W., Hood, D. C., & Anderassen, P. (1985). I. Was the South Sea Bubble a random walk?. *Journal of Economic Behavior & Organization*, *6*(4), 323-329.

Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of political Economy*, *111*(6), 1183-1220.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R Journal*, *8*(1), 289–317.

Scrucca, L., & Raftery, A. E. (2015). Improved initialisation of model-based clustering using Gaussian hierarchical partitions. *Advances in data analysis and classification, 9*(4), 447-460.

Sharma, S., & Escobari, D. (2018). Identifying price bubble periods in the energy sector. *Energy Economics*, *69*, 418-429.

Sohn, H., & Sornette, D. (2017). Bubbles as violations of efficient time-scales. *Institute for New Economic Thinking Working Paper Series*, (65).

Sornette, D. (2017). *Why stock markets crash: critical events in complex financial systems* (Vol. 49). Princeton University Press.

Swartz, N. D. (2014). Bursting the Bitcoin bubble: The case to regulate digital currency as a security or commodity. *Tul. J. Tech. & Intell. Prop.*, *17*, 319.

Thode Henry C. (2002). Testing for Normality. Statistics: textbooks and monographs

Thompson, E. A. (2007). The tulipmania: Fact or artifact?. *Public Choice*, *130*(1-2), 99-114.

Tirole, J. (1982). On the possibility of speculation under rational expectations. *Econometrica: Journal of the Econometric Society*, 1163-1181.

Venter, J. R., Griebenow, G., & De Jongh, P. J. (2005). Nig-Garch models based on open, close, high and low prices: theory and methods. South African statistical journal, 39(2), 79-101.

Von Hammerstein, E. A., Lütkebohmert, E., Rüschendorf, L., & Wolf, V. (2014). Optimality of payoffs in Lévy models. International Journal of Theoretical and Applied Finance, 17(06), 1450041.

West, K. D. (1987). A specification test for speculative bubbles. *The Quarterly Journal of Economics*, *102*(3), 553-580.

White, E. N. (1990). The stock market boom and crash of 1929 revisited. *Journal of Economic perspectives*, *4*(2), 67-83.

Wilhelmsson, A. (2009). Value at Risk with time varying variance, skewness and kurtosis—the NIG-ACD model. The Econometrics Journal, 12(1), 82-104.

Yamazaki, A. (2014). Pricing average options under time changed Lévy processes. Review of Derivatives Research, 17(1), 79-111.

Yan, W., Woodard, R., & Sornette, D. (2012). Diagnosis and prediction of rebounds in financial markets. *Physica A: Statistical Mechanics and its Applications*, *391*(4), 1361-1380.

Yeoh, P. (2010). Causes of the global financial crisis: Learning from the competing insights. *International Journal of Disclosure and Governance, 7*(1), 42-69.

Zhang, Q., & Wu, B. Y. (2018). An Anatomy of Stock Market Bubbles. *Canadian Social Science*, *14*(8), 36-44.

Zhang, Q., Zhang, Q., & Sornette, D. (2016). Early warning signals of financial crises with multi-scale quantile regressions of Log-Periodic Power Law Singularities. *PloS one*, *11*(11).

Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017, June). An overview of blockchain technology: Architecture, consensus, and future trends. In *2017 IEEE international congress on big data (BigData congress)* (pp. 557-564). IEEE.

Zhou, W., Chen, J., & Huang, Y. (2019). Co-Citation Analysis and Burst Detection on Financial Bubbles with Scientometrics Approach. *Economic research-Ekonomska istraživanja*, *32*(1), 2310-2328.

PUBLISHED ARTICLES

Cerecedo Hernández, D., Franco-Ruiz, C. A., Contreras-Valdez, M. I., & Franco-Ruiz, J. A. (2019). Explosión en Activos Virtuales (Criptomonedas). *Revista mexicana de economía y finanzas*, *14*(4), 715-727.

Núñez, J. A., Contreras-Valdez, M. I., & Franco-Ruiz, C. A. (2019). Statistical analysis of bitcoin during explosive behavior periods. *Plos one*, *14*(3), e0213919.

CURRICULUM VITAE

Carlos Armando Franco Ruiz was born in Mexico City, Mexico, on February 10, 1993. He was accepted in the PhD program in August 2017.

Experience:

2017 – Present **Exto Capital Multi Family Office**

- Chief Executive Officer

Portfolio performance, asset allocation, security selection, and financial instruments development.

2017 – Present **J&A Celorios, S.C.**

- Associate in Forensic Services

Economic and financial analysis to prevent fraud, economic crime, and disputes. Saving businesses from both financial and reputational damage.

2015 – 2017 **Allianz Mexico**

- Investment analyst

Investment analysis, and trust portfolio manager.

Education:

2011 – 2015 Bachelor's Degree in Economics and Finance **ITESM CCM**

Languages:

English

Microsoft Office

R

This document was typed in using Microsoft Word by Carlos Armando Franco Ruiz.