Louvain School of Management

Analysis of Stock Bubbles: Detection, Impact on Market Stability, and Portfolios' Reactions

Author: Guillaume BONINSEGNA
Supervisor: Bertrand CANDELON
Academic year 2023-2024
Dissertation for the Master [120] in Business Engineering with a focus in Finance
Daytime schedule

During the preparation of this master’s thesis, the author utilized ChatGPT from OpenAI for the following purpose:

1. To proofread the master’s thesis and correct language errors.

2. To debug the R code and improve the process.

After using ChatGPT from OpenAI, the author diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master’s thesis reflects our original work, augmented by the responsible use of AI.

Abstract :

Stock bubbles have been the triggering event of many crises. Therefore, it is crucial to understand these events and prevent them. The objectives of this thesis are to learn how to identify these phenomena, understand their consequences on market stability, and investigate their effect on investment portfolios. This paper is guided by the following research question: “Which strategies help in discerning stock bubbles, and to what extent do these phenomena impact overall market stability and the performance of investment portfolios?”. For this purpose, various methods were used. Firstly, regarding the first objective, we used various bubbles detection algorithms on R such as the GSADF and structural break tests. We also looked at deviations from historical trends and fundamental valuations, and differences in return distribution. Using the dot-com bubble as a practical case study and specific companies like Microsoft, we gained insight into how to cross results from different tests to identify bubbles. In addition, we also performed tests to assess the duration and intensity of the dot-com bubble on Microsoft and looked at factors that could have influenced the bubble. We gained valuable insights showing a link between the NASDAQ index around the dot-com bubble and the USA’s GDP, the consumer opinion index, and the consumer price index. Then, regarding the second objective, through GARCH analysis, we observed a positive correlation between the volatility of Microsoft, the NASDAQ and the S&P 500 around the bubble. We also gained insight into the positive correlations that exist between different asset classes, emphasizing the potential propagation of these events. Finally, regarding the third objective, we constructed four different portfolio types: an aggressive one, a conservative one, a balanced one, and an international diversified one. We then backtested them between 1995 and 2005 and observed that the balanced one had the best risk-return, in line with the benefits of diversification seen in the literature review. We compared this portfolio with a risk-free rate through the use of the Sharpe ratio, and found that the balanced portfolio was offering a better risk-return tradeoff.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm

Acknowledgements

This master's thesis closes my 5 years at UCLouvain as a business engineering student. I would like to thank all the professors and students who accompanied me on the journey. I would like to thank especially Professor Candelon, for his support and advice throughout the drafting of this thesis. And finally, I would like to thank my family and friends for their support.

Table of Contents

1. Introduction 1

1.1 Background 1

1.2 Research Objectives 1

1.3 Research Question 2

1.4 Methodology 2

1.5 Scope and Limitations 2

1.6 Thesis Structure 3

2. Literature Review 3

2.1 Stock Bubbles Overview 3

2.2 Fundamental Analysis of Stocks 4

2.3 Econometric Methods for Detecting Stock Bubbles 4

2.4 Implication of Stock Bubbles for Market Stability 5

2.5 Theories on Market Efficiency and Investor Behavior 5

2.6 Risk Management Strategies for Individual Portfolios 6

3. Identification and Analysis of Stock Bubbles 6

3.1 Identification of Historical Stock Bubble Episodes 7

3.1.1 Bubble Detection Algorithms 7

a) Data Collection 8

b) Stationarity Tests 8

ba) ADF Test on Raw Data 8

bb) PP Test on Raw Data 9

bc) KPSS Test on Raw data 9

bd) ADF Test on Differenced Data 10

c) GSADF Test 11

d) Structural Break Test on Raw Data 12

e) BDS Test 12

f) Analysis, Interpretation and Limitations 13
3.1.2 Deviations from Historical Trends 13
3.1.3 Deviations from Fundamental Valuations 14
a) Analysis of Price-to-Earnings (P/E) Ratio 14
b) Analysis of Dividend Yield 16
3.1.4 Comparative Analysis of Log Returns Distributions 18
3.2 Characteristics and Factors Contributing to Stock Bubbles 20
3.2.1 Duration and Intensity of Bubbles 20
3.2.2 Factors Contributing to Bubbles' Formation 20
3.2.3 Analysis of NASDAQ Composite Index and Macroeconomic Indicators 21
4. Impact of Stock Bubbles on Market Stability 23
4.1 Volatility Analysis 23
4.1.1 Implications of Volatility and Broader Market Impact 23
4.1.2 GARCH Model and Correlation Analysis 23
a) Day-to-Day Changes 24
b) Daily Return 25
4.2 Propagation of Stock Bubbles and Interdependence between Asset Classes 26
5. Portfolios' Responses 28
5.1 Portfolios Construction 28
5.1.1 Recall on Diversification 28
5.1.2 Portfolio Types 29
5.1.3 Asset Allocation 30
a) Aggressive Growth Portfolio 30
b) Balanced Portfolio 30
c) Conservative Portfolio 30
d) Diversified International Portfolio 30
5.2 Back-testing & Analysis 31

5.2.1 Back-testing 31
5.2.2 Comparison with a Benchmark: S&P 500 32
5.2.3 Comparison with a Risk-Free Investment / Sharpe Ratio 33
5.2.4 Conclusion and Limitations 34
6. Conclusion 34
7. Bibliography 36
8. Appendices 40
Appendix 1: Bubble Detection algorithm: R Script (3.1.1) 40
Appendix 2: Deviations from Historical Trends: R Script (3.1.2) 49
Appendix 3: Evolution of the P/E ratio of the S&P 500 between 1995 and 2005 (3.1.3) ... 51
Appendix 4: Comparative Analysis of Log Returns Distributions: R Script (3.1.4) 51
Appendix 5: Duration and Intensity of Bubbles: R Script (3.2.1) 55
Appendix 6: LM Model NASDAQ and Macroeconomic Indicators: R Script (3.2.3) 56
Appendix 7: GARCH Models & Correlation Analysis: R Script (4.1.2) 61
Appendix 8: Correlation Matrix between Asset Classes: R Script (4.2.1) 66
Appendix 9: Evolution of International Indexes and Correlation: R script (4.2.1) 68
Appendix 10: Gold price vs S&P 500 Adjusted Close Price: R Script (5.1.1) 71
Appendix 11: Table Computation of Sharpe Ratio, Balanced Portfolio 1995-2005 (5.2.5) 73

1. Introduction

1.1 Background

Stock bubbles have been the triggering event of many crises. Some notorious examples include the dot-com bubble, the housing-market bubble of 2008, or the Tulip Mania. Therefore, it is crucial for economists and individual investors to understand these phenomena and learn how to detect them. Even in today's world, where multiple financial studies have been done in this field, it is still difficult to accurately detect these phenomena. Stock bubbles are episodes when the price of an asset is above its real value, mainly driven by investors' enthusiasm. When the bubble bursts, the consequences can be terrifying. Indeed, on an individual level, while some techniques exist to diversify risk in individuals' portfolios, stock bubbles can erode all investors' profits. On a macroeconomic level, the consequences of a bubble burst can last for a long time before the economy returns to normal.

1.2 Research Objectives

In this context, the focus of this master's thesis is to explore the curiosity of stock bubbles, by gaining insight on their detection, their characteristics, and their broader implications for investors and the whole economy.

The first objective is to learn how to identify and understand these phenomena. For this purpose, multiple methods were used such as the use of bubble detection algorithms, observation of deviations from historical trends and key characteristics, and changes in return distributions. Also, some analyses were done to gain insight into the key characteristics of bubbles. The dot-com bubble mainly serves as a case in point.

Understanding the consequences of these events is equally critical. Indeed, stock bubbles have significant implications on markets and global financial stability, potentially even leading to crises. This highlights the importance of knowing the implications of these events and their propagation. Various methods were used for this purpose such as volatility analysis, correlation studies, and evaluations of broader market impacts and propagation.

Finally, the last objective investigates the impact of a bubble burst on investment portfolios. This focuses on the importance of prudent portfolio management and strategies such as diversification, asset allocation, and portfolio rebalancing. This was analyzed through a practical case study (the dot-com bubble), back-testing, and comparisons with established benchmarks.

1.3 Research Question

The research objectives of this work led to the following research question:

"Which strategies help in discerning stock bubbles, and to what extent do these phenomena impact overall market stability and the performance of investment portfolios?"

1.4 Methodology

This section outlines the methodology used for the various tests used in this paper. The first step is data collection, which was done by sourcing financial data, historical stock market information, and scientific literature related to stock bubbles. Then, this thesis employed various statistical techniques for quantitative data analysis. Some techniques included deviations from historical trends and bubble detection algorithms for the examination of the presence and intensity of stock bubbles. It also included the examination of stock valuation metrics over time, like the Price-to-Earnings ratio or the dividend yield, to study their relationship with stock prices and market trends, and to see potential deviations around the period of a bubble. In addition, an example of GARCH model analysis was performed to assess volatility in the market and a correlation matrix between various asset classes was made. Lastly, hypothetical portfolios were made, and their performance analyzed around the period of a bubble. All the quantitative analysis and graphs were performed on R or on Excel, and the scripts can be found in Annexes.

1.5 Scope and Limitations

The thesis focuses on stock bubbles detection, impact on market stability and on investment portfolios with historical case studies. This work is principally focused on the period around the dot-com bubble. This choice is justified by the fact that the dot-com bubble is the most recent significant bubble in the American stock market. Using more recent US housing bubble or Chinese stock bubble for the analysis would not have been relevant because the focus of this thesis is on stocks and major financial crises induced by them. While every effort has been made to ensure the robustness of this study, it is important to acknowledge potential limitations such as data availability, generalizability of findings, time constraints, complexities, and external factors in establishing causality in financial markets.

1.6 Thesis Structure

This paper begins with a literature review in Section 2, which explores the phenomena of stock bubbles, the ways to detect them through fundamental analysis and econometrics, their implications for the market stability, theories on the efficient market hypothesis and investors' behavior, and risk management strategies to mitigate their effects on portfolios. Then, Section 3 delves into the identification and analysis of stock bubbles through bubble detection algorithms, analysis of deviations from historical trends and fundamental valuations, comparative analysis of returns distributions, and characteristics and factors contributing to stock bubbles. Thereafter comes Section 4, which investigates the impact of stock bubbles on market stability through volatility analysis and analysis of propagation to other asset classes. It is followed by Section 5, which explores the portfolios' responses by constructing hypothetical portfolios and backtesting them around the period of the dot-com bubble to see which ones perform the best. Finally, Section 6 concludes this paper and Section 7 highlights the relevant sources used for this work, while Section 8 is composed of the appendices.

2. Literature Review

2.1 Stock Bubbles Overview

Speculative bubbles could be defined as situations in which an asset has a price temporarily higher than the estimated real value. These events are often driven by investors' enthusiasm rather than fundamental analysis (Shiller, 2000). Some notorious examples include the South Sea Bubble, the Tulip Mania, the US housing bubble, or the dot com bubble (which will get a special attention in this thesis). These examples have all in common that the bubble burst led to a crisis. In addition, bubbles involve a capital misallocation across assets and sectors, leading to a decrease of welfare (Miao, 2014). This poses significant risks to the economy's stability.

Many authors searched for ways to detect these phenomena accurately. P.C.B. Phillips and Jörg Breitung stand as relevant examples, and some of their papers and econometrics methods are used in this thesis regarding bubbles detection.

Shiller (2000) emphasized the fact that bubbles are driven by irrational exuberance and market sentiment from investors. He argued that even a correctly applied fundamental analysis cannot be enough to reduce the number of stock bubbles.

2.2 Fundamental Analysis of Stocks

Fundamental analysis is a fundamental concept in finance and investment. It is a method widely used by investors to determine the value of a stock. It is relevant regarding the concept of stock bubbles, as it can help detect when an asset deviates from its intrinsic value. To evaluate a company, it concentrates on criteria such as financial performance, industry conditions and overall economic indicators (Graham & Dodd, 2008). Some famous books like “Security Analysis” (2008) or “The Intelligent Investor” (2003) from Graham and Dodd and Graham alone respectively, provide some kind of guidelines regarding how to evaluate investment opportunities. However, this concept has evolved since then, and now other factors are being considered such as earnings per share, free cash flow, return on equity, future growth, and industry benchmarks (Damodaran, 2012).

However, the use of fundamental analysis to detect bubbles is controversial. On one hand, some authors like Graham and Dodd (2008) argue that it could help detect overvalued stocks, which potentially indicate bubbles. On the other hand, other authors like Shiller (2000), argue that the unpredictable nature of investor behavior and market sentiment make it difficult for fundamental analysis alone to predict bubbles accurately.

From this discussion, a consensus could emerge that it is important to understand both fundamental analysis and behavioral finance to be able to detect stock bubbles.

2.3 Econometric Methods for Detecting Stock Bubbles

In addition to fundamental analysis, other methods exist in today’s world to detect stock bubbles. For this purpose, econometrics and statistics play a crucial role. Various advanced econometric tests exist including the notable ones introduced in Phillips, Shi, and Yu’s works (2011, 2015).

One of their key findings is the Generalized Sup ADF (GSADF) test, which is an extension of the Augmented Dickey-Fuller (ADF) test, a unit root test (Phillips et al., 2015). A unit root test examines if a time series follows a stochastic trend (potential bubble behavior) or a deterministic trend (behavior aligned with economic fundamentals). This test uses multiple

ADF tests with a rolling window over the sample. The goal is to find explosive movements in asset prices, potentially indicating bubbles.

Other tools exist such as the Chow Test which is a classical method for detecting structural breaks (Hansen, 2001), the Breitung's Rank Test which is also a non-parametric rank test for unit roots (Breitung, 2002), or the CUSUM (Cumulative Sum) test introduced by Page (1954) and extend by Brown, Durbin, and Evans (1975), which is used to detect structural breaks based on the sum of residuals.

2.4 Implication of Stock Bubbles for Market Stability

As mentioned in previous sections, stock bubbles can lead to market instability and crises. The capital misallocation linked with the overpricing of assets impacts this stability and destabilizes the market. For example, the 2008 global financial crisis, started by a bubble in the housing market, led to a huge economic recession (Schwartz, 2009). Another factor potentially increasing the likelihood of crisis after a bubble burst is what Geanakoplos (2010) called a "deleveraging" phase. This is when people seek to reduce their level of debt. Panic can further exacerbate a financial downturn.

In addition, Ali, Jan, and Jehanzeb (2015) highlight that investors' sentiments are influenced by fluctuations in the market and that investment decisions before and after a crash are significantly different. This means that the potential recession due to a bubble burst could even be exacerbated through more prudent behaviors from economic agents. Also, Tversky and Kahneman (1992), in their paper about cumulative prospect theory, assume that for risk averse people, losses are more impactful than gain on utility.

Finally, most authors agree on the fact that there is a relation between stock bubbles and volatility. Cuñado, Gil-Alana, and De Gracia (2009) found that, in accord with other authors, market volatility is higher during times of bear markets than in bull markets. It confirms that bubbles could negatively impact market stability when they burst.

2.5 Theories on Market Efficiency and Investor Behavior

According to Fama (1970), an "efficient" market is a market in which asset prices always "fully reflect" available information and provide accurate signals for investment and resource allocation. However, Shiller (2002) doesn't think that markets are efficient, and stock bubbles are a proof. Indeed, these phenomena highlight cases in which prices don't reflect intrinsic values, and thus are not optimal. Studies on behavioral finance often bring another view regarding the efficient market hypothesis. Shiller (2000), in his book "Irrational Exuberance", highlighted that investors' behavior is a key factor to the mispricing of assets and the forming of stock bubbles, due to some biases they have such as herd mentality or overconfidence. These theories are relevant for this paper as they help understand factors leading to the creation of stock bubbles.

2.6 Risk Management Strategies for Individual Portfolios

Knowing the terrible outcome of a bubble burst, it is crucial for investors and economic agents to develop risk mitigation strategies. Various methods exist such as diversification, hedging, stop-loss orders, or stress testing.

The most famous one is probably diversification, "don't put all your eggs in one basket", as recommended by the Modern Portfolio Theory of Markowitz (1952). This is one of the key risk mitigation strategies used by investors. It is closely linked with asset allocation, which plays a pivotal role in portfolio management. Ibbotson and Kaplan (2000) highlighted in their study that more than 90% of the variability of a fund's returns over time could potentially be attributed to asset allocation.

Another important approach regarding prudent portfolio management is the use of regular rebalancing. This means adjusting the asset allocation periodically, by selling those that have appreciated beyond their estimated value and buying those considered undervalued. For example, Graham (2003) suggested that an investor should often adjust his portfolio to have between 25% and 75% composed of common stocks and an inverse range between 75% and 25% in bonds.

In addition to other methods applied to reduce variance in portfolios, there is also another kind of strategy regarding stock bubbles: learning how to potentially recognize them. A lot of authors and researchers already worked on it and identified some patterns. But for this work, the primary sources used for this purpose will be from the papers of Phillips and his co-authors (2011, 2015). Being able to recognize potential bubbles could help investors make more informed decisions and safeguard their investments by adjusting asset allocation in their portfolios.

3. Identification and Analysis of Stock Bubbles

In this section, we focused on the identification and analysis of stock bubbles, with a historical case in focus: the dot-com bubble. This section investigates the identification of these events, their characteristics, and factors contributing to them. We concentrated on major tech companies and market indexes from 1995 to 2005, because it captures the formation and bursting of the dot-com bubble. Indeed, according to the finding of Phillips, Wu, and Yu (2011), the explosive trend of exuberance started in 1995 and ended in about 2000-2001. This period, especially the late 1990s, was characterized by a surge in investment in companies related to the internet (dot-coms). In 2000, the bubble burst when people saw that these companies were overvalued, leading to a fast decline in the market.

3.1 Identification of Historical Stock Bubble Episodes

Identification is the first step and probably the most critical one when relating to bubbles. For this purpose, various methods were used such as some bubble detection algorithms, and analysis of deviations from historical trends and fundamental valuations. Phillips, Shi, and Yu (2015) define financial bubbles following this equation:

$$P_t = \sum_{t=0}^{\infty} \left(\frac{1}{1+r_f}\right)^t E_t(D_{t+i} + U_{t+i}) + B_t,$$

where P_t is the price of the asset after-dividend, D_t is the cash flow coming from the asset, r_f is the risk-free interest rate, U_t is the unobservable fundamentals, and B_t is the bubble. $P_t^f = P_t - B_t$ is considered as the market fundamental. If there is no bubble, $B_t = 0$. (Phillips et al., 2015) More basically, bubbles could be defined as:

$$B_t = P_t - F_t,$$

where F_t is the price estimated through fundamental analysis. However, as noted in the literature review, estimating an asset value through fundamental analysis may be challenging, knowing that it involves a lot of criteria.

3.1.1 Bubble Detection Algorithms

We applied different tests on Microsoft’s stock price data from 1995 to 2005. The results are displayed in the following sections. Figure 1 represents Microsoft's daily adjusted closing price during this period. We can observe an important increase until the year 2000, followed by a sharp decline, potentially indicating the bubble.

Figure 1: Microsoft Adjusted Close Price, 1995-2005

All the following tests were performed on the R programming language and the scripts used can be found on Appendix 1.

a) Data Collection

The historical data used were retrieved using the "getSymbols" function from the "quantmod" package in R. The data come from Yahoo Finance (2023) and are dated during the period from January 1, 1995, to December 31, 2005.

b) Stationarity Tests

Unit root tests were performed to assess the stationarity of data. This is important because we used a method for identifying structural breaks in a following section, which assumes that data are stationary. Stationarity means that the statistical properties (such as mean or variance) of the time series data do not change over time. Conversely, non-stationarity could indicate different market behavior. To assess it, we used the Augmented Dickey-Fuller (ADF) test, the Phillips-Perron (PP) test, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test.

ba) ADF Test on Raw Data

We employed the Augmented Dickey-Fuller test to test the null hypothesis that the time series data has a unit root against the alternative hypothesis that the time series data is stationary or trend-stationary (Dickey & Fuller, 1979). We used the "urca" package on R to perform this test.

The equation that represents this test is:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-1} + \varepsilon_t,$$

where Δy_t is the difference in the time series, α is a constant, β is the time trend component, γ is the coefficient on the lagged level of the series, δ_i are the coefficients on the lagged differences, and ε_t is the error term (Phillips, Shi & Yu, 2015). The null hypothesis H_0 is the presence of a unit root, $\gamma = 0$, and the alternative hypothesis H_a is the non-presence of a unit root, $\gamma < 0$.

The ADF test showed a test statistic value of -2.0172, which is more than the critical value at 5% level (-2.86). This indicates that it failed to reject the null hypothesis of a unit root. Therefore, the Microsoft stock price series from 1995 to 2005 is supposed to not be stationary according to this test.

bb) PP Test on Raw Data

The Phillips-Perron test is another unit root test which controls for autocorrelation and heteroskedasticity in the error term. This makes it robust to a wider array of circumstances in finance and it is useful combined with the ADF test. We also used the "urca" package in R to perform the PP test. This test could be expressed in a simplified version as:

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

where y_t is the time series as in the ADF test, ρ is the autoregressive coefficient, and ε_t is the error term. The null hypothesis H_0 is that $\rho = 1$, indicating that there is a unit root. (Phillips & Perron, 1988)

The PP test resulted in a test statistic (Z-tau) value of -2.035, which is more than the critical value at 5% level (-2.86). Thus, it again failed to reject the null hypothesis of a unit root, providing further evidence and solidifying our conclusion that the series is not stationary.

Considering these results, we can conclude that the Microsoft stock price series from 1995 to 2005 is non-stationary according to both the ADF and PP tests.

bc) KPSS Test on Raw data

In addition to the two previous tests which were focusing on unit root, the KPSS test completes these methods by adding a different perspective on stationarity. The Kwiatkowski-Phillips-Schmidt-Shin test serves as an alternative to the ADF test for checking the stationarity of a time series. Unlike the ADF test where the null hypothesis presumes the presence of a unit root (non-stationary series), the KPSS test assumes the series is stationary under the null hypothesis. (Kwiatkowski, Phillips, Schmidt, & Shin, 1992) The equation of this test is:

$$y_t = r_t + \beta t + \varepsilon_t,$$

where r_t is a random walk, βt is a deterministic trend, and ε_t is the error term. (Kwiatkowski et al., 1992)

The computed KPSS test statistic is 5.0697. It considerably exceeds the critical values across all specified significance levels (0.119 at 10%, 0.146 at 5%, 0.176 at 2.5% & 0.216 at 1%). This confirms the rejection of the null hypothesis in favor of the alternative, indicating that the time series is not stationary, and has potentially a unit root.

Therefore, combining these three tests, we can deduce that Microsoft's adjusted stock prices are not stationary, which is not unusual for this type of financial time series.

bd) ADF Test on Differenced Data

Given the non-stationarity of the raw data, we tried to stationarize it by considering the differences. For this purpose, we applied first-order differentiation to the series and visualized it on the following graph, which shows day-to-day changes in stock prices. This section is only for contextual analysis, to give a fuller picture of the situation.

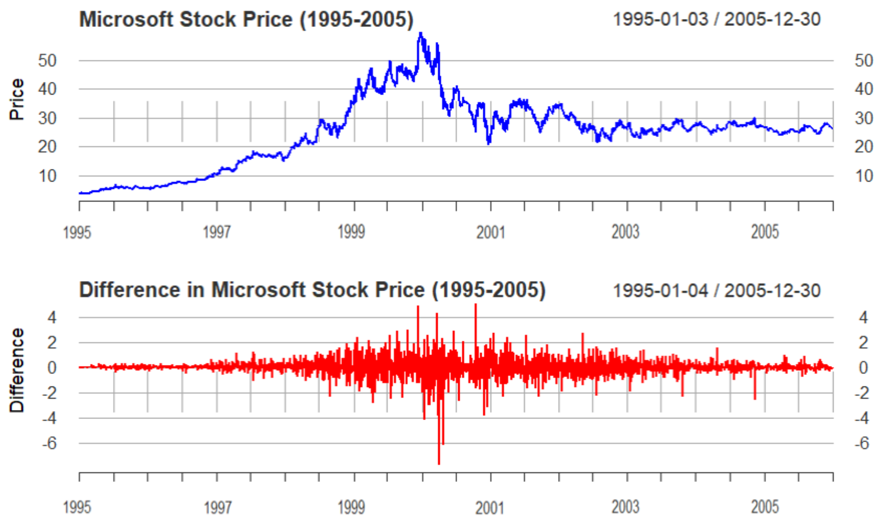

Figure 2: Difference in MSFT Stock Price, 1995-2005

From this visualization of Microsoft's stock prices from 1995 to 2005 and its daily changes (Figure 2), we can notice spikes in the differenced plot (the second one) around the year 2000. This represents day-to-day changes in this asset's price, thus a higher volatility than usual. This can likely be attributed to the influence of the dot-com bubble.

After differencing, we applied once again an ADF test, but this time on the differenced data. The resulting test statistic was a considerably more negative value of -36.5969. This significant shift in the test statistic confirms that differencing successfully rendered the time series stationary. To further illustrate, when compared to the critical value of -2.86 at the 5%

significance level, this value decisively rejects the null hypothesis of a unit root, indicating stationarity in the differenced series.

c) GSADF Test

The Generalized Sup Augmented Dickey-Fuller (GSADF) test, developed by Phillips, Wu, and Yu (2015), is an extension to the ADF test and other traditional unit root tests. It employs a rolling window approach (30 days in this analysis), making it more accurate.

Following the methodology outlined by Phillips, Shi, and Yu (2015), the GSADF test was applied to the daily closing prices of MSFT stock from January 1, 1995, to December 31, 2005. To stabilize variance, the data were log-transformed: for a time series Y_t, $X_t = \log{(Y_t)}$. The GSADF test then evaluated the null hypothesis of a unit root against the alternative of explosive behavior (indicating a bubble) within a rolling window of fixed size (30 days).

$$GSADF_w(t) = \max_{w \le r \le t} ADF(r, t),$$

where $ADF(r, t)$ is the traditional ADF test calculated for the sample from r to t.

The R scripts used can be found in Appendix 1 as for the other tests.

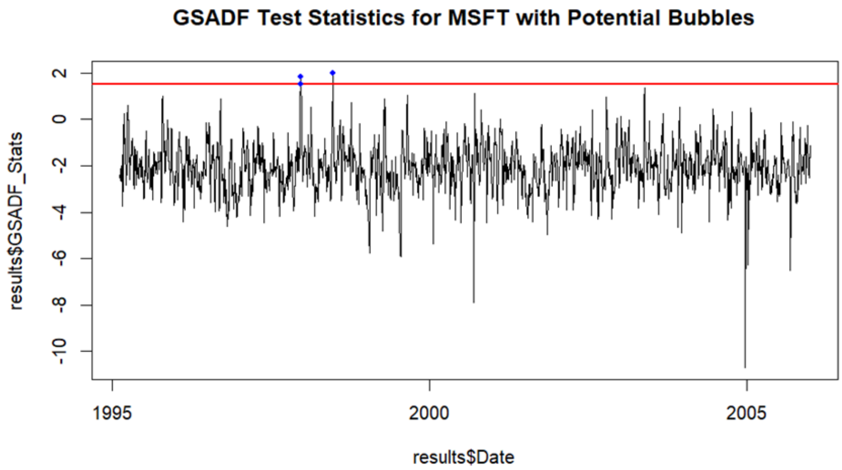

Figure 3: GSADF Test for MSFT, 1995-2005

The red line in Figure 3 represents a threshold based on the 95th percentile of the GSADF statistics and the blue points indicate three moments where the GSADF statistics were significantly higher than usual, above this threshold, potentially highlighting the bubble. The exact dates at these points are December 23 and 24, 1997, and June 24, 1998. These dates could be linked with corporate events but also to the tech industry’s expansion in the late 1990s, around the dot-com bubble.

d) Structural Break Test on Raw Data

The structural break test was performed using the "breakpoints" function from the "strucchange" package on R. Given the interest in identifying significant shifts in the level of Microsoft's stock price, we opted to apply the structural break test to the raw, non-stationary data. Because while differencing can make a series stationary, it also transforms the nature of the data, emphasizing changes rather than levels. In this context, the raw data were more relevant regarding the research questions. Note that one could perform the analysis on differenced data with the same method used in raw data (but we have made this test for an informative purpose and have found no clear structural breaks). We identified four breakpoints at specific dates: April 9, 1997, November 30, 1998, July 25, 2000, March 28, 2002. Figure 4 is a visual representation.

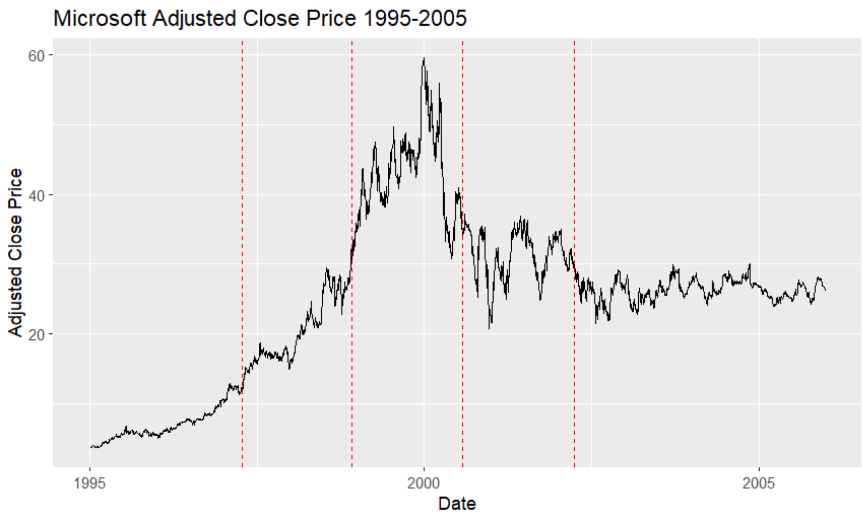

Figure 4: MSFT Adjusted Close Price with Breakpoint, 1995-2005

e) BDS Test

The BDS (Brock, Dechert & Scheinkman) test looks at nonlinearity and dependence in time series data. Therefore, this test adds new characteristics of price movement in consideration to detect bubbles. This test, investigated by Brock, Dechert, and Scheinkman in 1996, has a null hypothesis that states that the residuals from a time series are independent and identically distributed (i.i.d). This is in line with the efficient market hypothesis seen in literature review, which suggests that stock prices follow a random walk pattern.

In contrast to this hypothesis, the test returned low p-values across all dimensions, suggesting the presence of non-random patterns in Microsoft's stock prices during the period from 1995 to 2005. This result challenges the notion of an efficient market and could potentially be linked to a bubble, as values deviate from their fundamental valuations.

f) Analysis, Interpretation and Limitations

Firstly, the non-stationarity tests (ADF, PP and KPSS tests) we performed all highlighted the presence of a unit root and non-stationarity, indicating important changes in Microsoft's stock prices. This could be a clue indicating the dot-com bubble. In addition, the GSADF test also indicated a potential signal of a bubble, especially in 1997-1998. Then, the structural break test identified four dates, probably attributed to specific corporate or economic events. Looking at these dates, we cannot deduce a bubble from this test. Finally, the BDS test detected a non-linearity in the data, challenging the efficient market hypothesis, and further providing a pattern regarding a bubble.

These tests are not enough to accurately provide evidence of the existence of a bubble, but the patterns observed could be indicators to make further research. We should also note that interest in only one stock could be biased when looking around a big bubble such as the dot-com one and a single stock could be influenced by multiple other factors.

3.1.2 Deviations from Historical Trends

To further analyze patterns to detect stock bubbles, we compared Microsoft's stock prices with their historical averages, still around the dot-com bubble. For this purpose, we compared with a 90-day rolling average (Figure 5). We identified periods where the price significantly exceeds this range, potentially indicating the bubble.

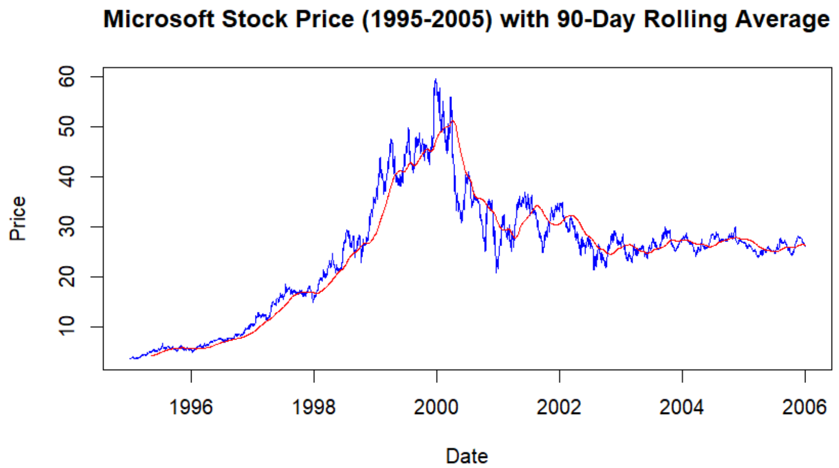

Figure 5: MSFT Stock price and 90-Day Rolling Average, 1995-2005

Then, we computed the differences (residuals) between the actual prices and the 90-day rolling average at each time (Figure 6).

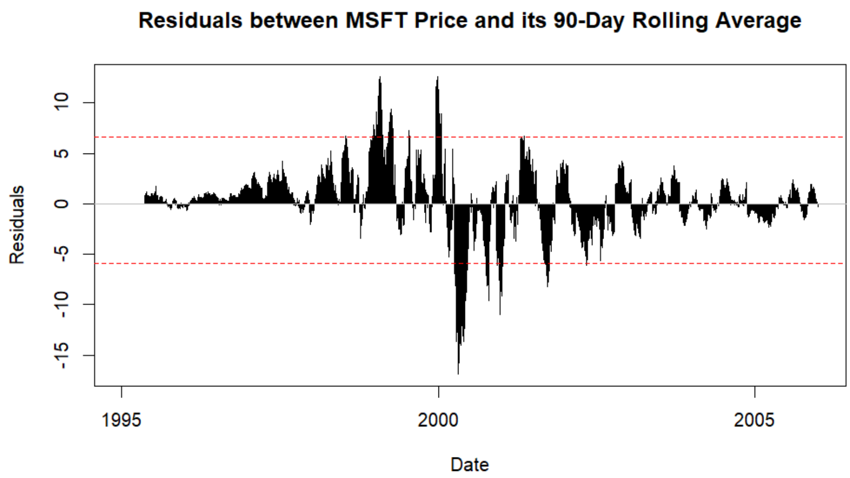

Figure 6: Residuals between MSFT Price and its 90-Day Rolling Average, 1995-2005

Positive residuals indicate moments when Microsoft's stock price was above its 90-day average, and conversely. We can observe large positive differences in 1999 and 2000, and then large negative differences during the year 2000. This is in line with the historical dot-com bubble, which burst in 2000. Note that we consider as large movements those above a threshold (red lines on the graph) of 2 standard deviations from the mean of the residuals (as it encapsulates approximately 95% of the data in a normal distribution). The maximum positive residual is on December 27, 1999, during the upward period of the bubble; and the maximum negative residual is on April 24, 2000, the year in which the bubble burst.

This method, combined with others, is a visual indicator to identify and understand the scale of a bubble.

3.1.3 Deviations from Fundamental Valuations

After having discussed the concept of fundamental analysis and its relation to stock bubbles in the literature review, we tried to analyze it practically. By doing it, we aimed at gaining insight into potential bubble indicators. For this purpose, we still focused on the period around the dot-com bubble (from 1995 to 2005) and looked at major tech companies such as Microsoft (MSFT), Adobe (ADBE) and Intel (INTC).

a) Analysis of Price-to-Earnings (P/E) Ratio

Price-to-Earnings ratio is a criterion commonly used in fundamental analysis to assess a stock value. It is calculated as follows:

$$\frac{P}{E}\,Ratio = \frac{Share\ Price}{Earnings\ per\ Share\ (EPS)},$$

where the share price is the current stock price and EPS represent the part of the company's profit for each share. EPS is calculated as:

$$EPS = \frac{(Net\ Income - Dividends\ on\ Preferred\ Stock)}{Weighted\ Average\ Number\ of\ Shares\ Outstanding}.$$

The data for the stock prices were retrieved from Yahoo Finance (2023) and the EPS from Ycharts.com (2023). We used the adjusted close prices, like in the other analysis, rather than close prices because it better reflects reality of the value of shares, accounting for actions such as dividends or stock splits. Concerning the EPS, we used diluted EPS rather than the basic one because it gives a more prudent view, as it assumes that all convertible securities are indeed converted.

Figure 7 is a graph of the evolution of the P/E ratios of Microsoft during the period from 1995 to 2005.

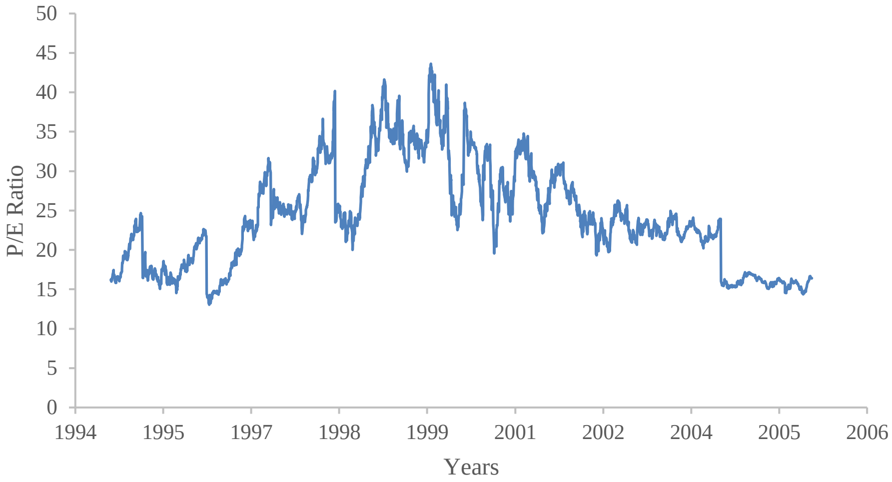

Figure 7: P/E Ratios Evolution of MSFT, 1995-2005

A high P/E ratio indicates that a company's stock price is high compared to its earnings, and thus potentially overvalued, and conversely. However, looking at its evolution for only one company is not relevant enough. To have a critical view, we must compare it to a benchmark, or to historical values. We chose to compare these ratios with the S&P 500 P/E ratios over the same period (retrieved from UpMyInterest.com (2023)). We didn't compare with companies from the same sector, in the NASDAQ, because it would have provided uninteresting results as this whole market experienced the dot-com bubble.

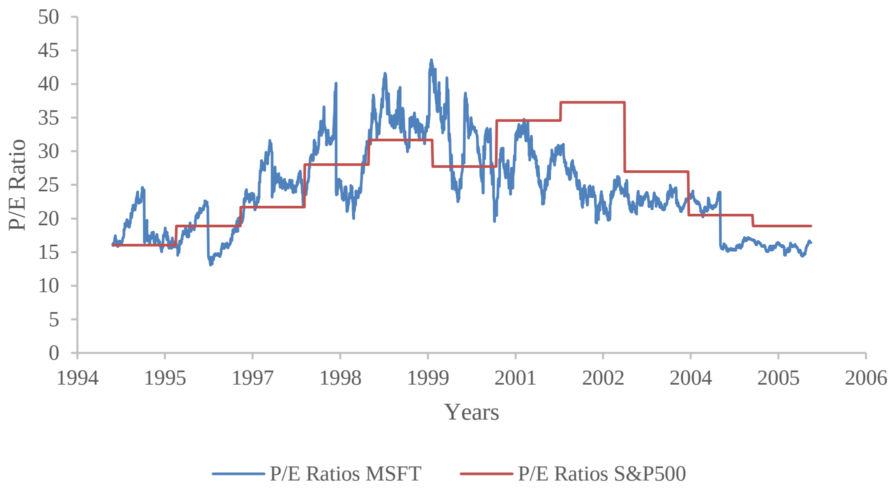

Figure 8: P/E Ratios MSFT vs P/E Ratios S&P 500, 1995-2005

We can observe that the P/E ratio of Microsoft was significantly higher than the broader market, the S&P 500, in 2000 (Figure 8). This could indicate an overvaluation, potentially driven by investors' expectations. The reverse trends from 2000 to 2002 probably indicate the burst of the bubble. We should note that the S&P 500 includes tech companies, which means that it was also impacted by the bubble, but with a lower effect due to the diversity of this index. However, we should keep in mind that the P/E ratio has some limitations such as that it doesn't account for factors like expected growth, companies with negative earnings, volatility, or debt.

b) Analysis of Dividend Yield

Dividend yield is another tool commonly used in fundamental analysis to assess a company's value. It indicates how much a company pays dividends per share each year. Like the P/E ratio, this metric is only an indicator that has to be considered in a broader view, with other criteria and comparisons with benchmarks. Dividend yield is a percentage and is calculated as follows:

$$Dividend\ Yield = \frac{Annual\ Dividends\ per\ Share}{Price\ per\ Share}.$$

Still focusing on the dot-com bubble, we noticed that Microsoft only started paying dividends in 2003, after the crisis, and Apple didn't pay dividends between 1996 and 2012, due to a difficult period. Therefore, we focused on two other tech companies: Intel (INTC) and Adobe (ADBE), which both have paid dividends during our period of interest. The raw data were retrieved from Yahoo finance (2023) and companies' official websites. We computed dividend yields over a larger period, from 1990 to 2010, to better see a potential change.

Figure 9 represents the evolution of the dividend yield of INTC.

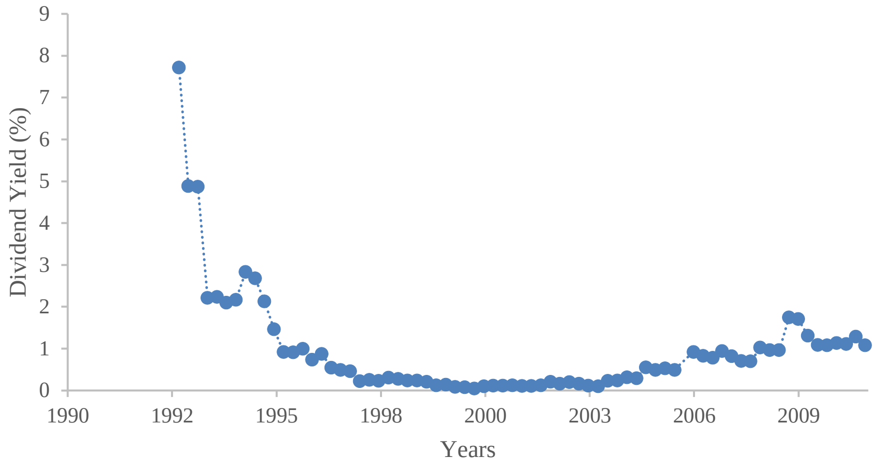

Figure 9: Dividend Yield of INTC, 1990-2010

It indicates a gradual decline until 2000, where their dividend yield was at its lowest value (0.46794%) the first of September, and then a rise. This trend could be linked to the dot-com bubble.

Similarly, for Adobe:

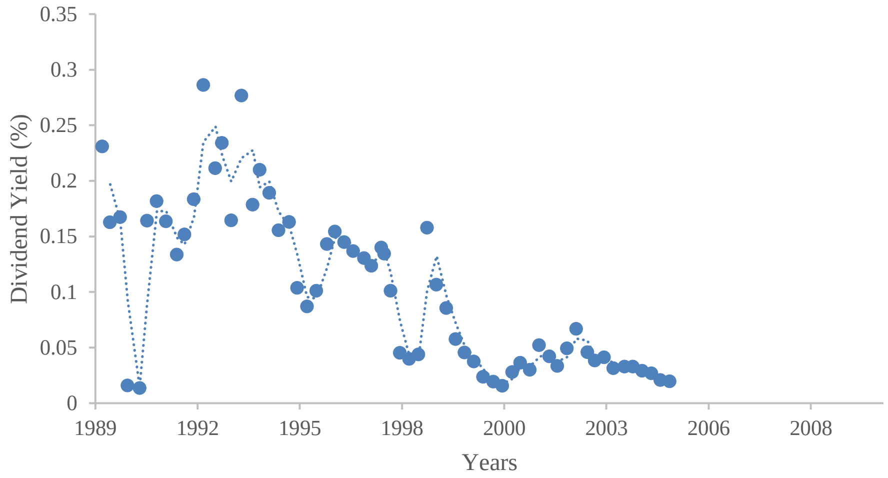

Figure 10: Dividend Yield of ADBE, 1990-2010

A similar pattern to that of INTC could be observed around 2000 (Figure 10).

Then, we made the same analysis for the S&P 500 (Figure 11).

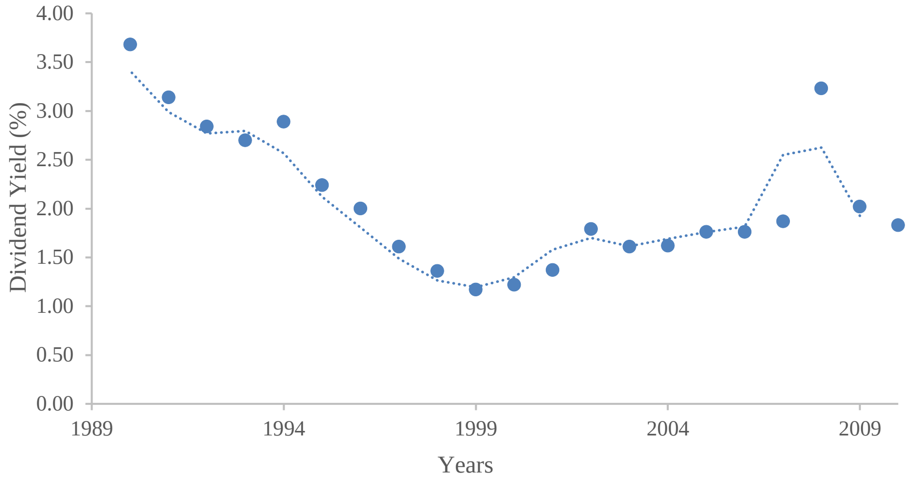

Figure 11: Dividend Yield of S&P 500, 1990-2010

The same pattern is again observable for this broader market, even if the values are significantly higher than for the two tech companies.

A high dividend yield indicates that companies return a large amount of profits to investors relative to stock prices, and conversely. A low dividend yield may indicate that companies use their profit to invest in growth rather than distributing it, or that the stock's price is too high relative to earnings. In our analysis, it could indicate that investors overestimated companies' growth and that the profits of companies were not as good as expected.

3.1.4 Comparative Analysis of Log Returns Distributions

This section investigates the different distributions of the log returns of stock prices over different periods. For this analysis, we compared the period from 1995 to 2005, capturing the dot-com bubble, and the period from 2006 to 2016, a slightly more stable period for tech companies after the bubble.

Firstly, we looked at MSFT.

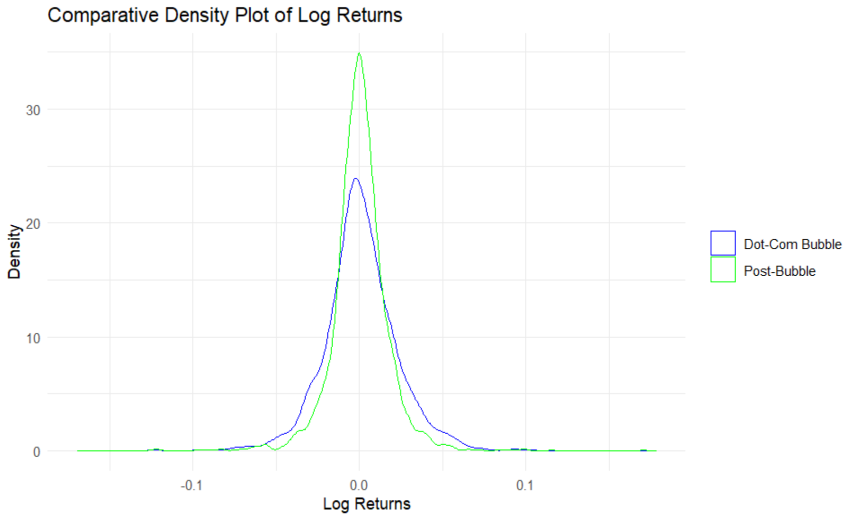

Figure 12: Comparative Density Plot of Log Returns MSFT, Dot-Com Bubble vs Post-Bubble

We can observe that the distribution post-bubble is more stable (Figure 12). We observe that the mean and standard deviation in the period around the bubble (0.00075 and 0.023) were significantly higher than in the post-bubble period (0.0004 and 0.017), indicating higher returns in average but also higher volatility. We also performed a Kolmogorov-Smirnov test to verify the differences between the distributions of the two periods and this test strongly rejected the null hypothesis of identical distributions, confirming that they behave differently.

We proceeded to the same analysis with INTC and got the same kind of observation, with a significantly higher mean and standard deviation during the dot-com bubble (Figure 13).

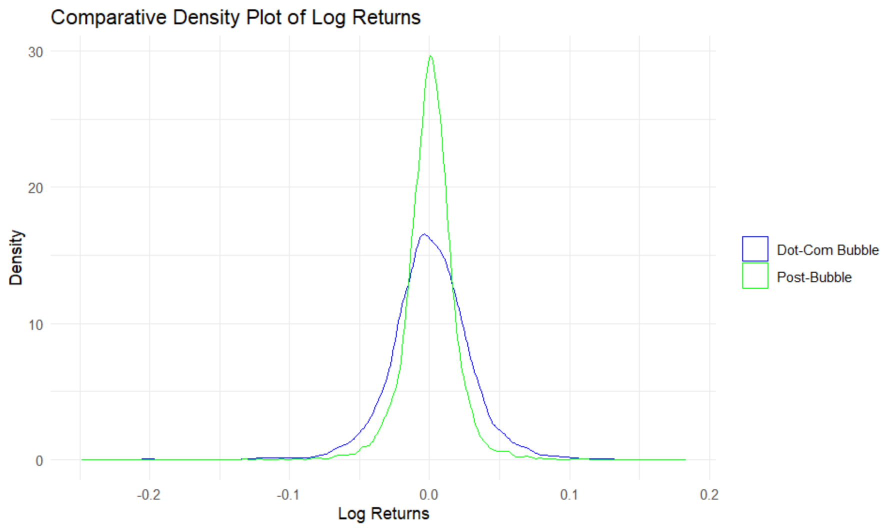

Figure 13: Comparative Density Plot of Log Returns INTC, Dot-Com Bubble vs Post-Bubble

Moving to the whole NASDAQ market through the ^IXIC index, we observe the same kind of results (Figure 14).

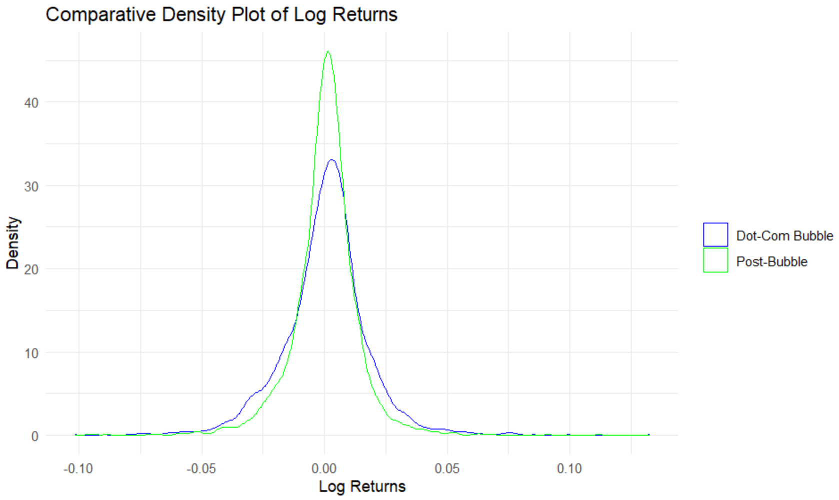

Figure 14: Comparative Density Plot of Log Returns ^IXIC, Dot-Com Bubble vs Post-Bubble

3.2 Characteristics and Factors Contributing to Stock Bubbles

3.2.1 Duration and Intensity of Bubbles

To measure the potential impact of bubbles, duration and intensity are two important criteria. Duration gives information about how long the inflated prices have lasted on the market, while intensity measures how much a price increased compared to the initial price. Focusing on MSFT during the dot-com bubble, we computed on R (scripts in Annexes) a duration of 437 days, and an intensity of 102.66%. For this purpose, the R script looked at the moment when the stock price significantly deviates from its average trends, and the moment when it has come back to normal. The results indicated a long period of inflated prices, more than one year, whereas the stock price more than doubled.

3.2.2 Factors Contributing to Bubbles' Formation

Firstly, the macroeconomic conditions can highly influence the formation of a stock bubble. These include excessive liquidity and low interest rates. Too much liquidity can lead to asset inflation, and low interest rates give opportunity to people to invest more than they usually could, by borrowing money. Okina, Shirakawa, and Shiratsuka (2001) highlight that it was the second factor that contributed the most to the bubble of the late 1980s in Japan. Another notorious example of the correlation between interest rate and bubble is the housing bubble of 2008. In addition, another macroeconomic factor that drives bubbles and their impacts is the interdependence between markets globally. These relations between global markets can amplify the bubble and propagate the potential crisis that arises from the burst. This effect of contagion is highlighted in the paper “Manias, Panics, and Crashes: A History of Financial Crises” of Kindleberger (2005).

Then, investors' behaviors such as herd mentality, over-optimism and confirmation bias can also influence the formation of bubbles. Herd mentality refers to an event or period when people act as followers of a group. (Bikhchandani & Sharma, 2001) In the stock market, this could mean selling or buying assets when others do. Therefore, it is very problematic regarding stock bubbles. Over-optimism and speculation are also at the roots of these phenomena. They relate to moments when investors overestimate the growth potential of stocks, underestimating the risks. This leads them to buy stocks rapidly to get profit, and thus inflating the bubble. (Shiller, 2000) Regarding confirmation bias, this means emphasizing information that confirms our thought and potentially ignoring the others. This could also emphasize the forming of bubbles.

Finally, the regulatory environment also contributes to the formation of bubbles. Okina and his co-authors (2001) highlight the progressive financial deregulation in Japan in the late 1980s, leading to aggressive bank behavior and then the bubble. This deregulation, as well as the new financial tools available, gives easier access to the market. This easy access for everyone can boost bubbles if investors are not experienced enough. In addition, a lack of, or an inadequate, regulatory framework can also foster bubble formations.

3.2.3 Analysis of NASDAQ Composite Index and Macroeconomic Indicators

This section investigates the relationships between the NASDAQ and some macroeconomic indicators around the period of the dot-com bubble. For this purpose, we performed a linear regression model formulated as follows:

$$IXIC.Adjusted = \beta_0 + \beta_1 USALORSGPNOSTSAM + \beta_2 UNRATE + \beta_3 UMCSENT + \beta_4 MVMTD027MNFRBDAL + \beta_5 CSCICP03USM665S + \beta_6 CPIAUCSL + \beta_7 FEDFUNDS + \varepsilon,$$

where β_0 is the intercept, β_1, β_2, β_3, β_4, β_5, β_6 and β_7 are the coefficients that represent the impact of each variable, and ε is the error term. $IXIC.Adjusted$ represents an index (^IXIC) of adjusted closing prices of the NASDAQ retrieved from Yahoo Finance (2023). The other variables: consumer sentiment ($UMCSENT$), unemployment rate ($UNRATE$), consumer price index ($CPIAUCSL$), consumer opinion survey ($CSCIP03USM665S$), market value of marketable treasury debt ($MVMTD027MNFRBDAL$), gross domestic product for United

States ($USALORSGPNOSTAM$), and federal funds effective rate ($FEDFUNDS$), were retrieved from the Federal Reserve Economic Data (FRED) (2024). The period of interest is between 1995 and 2005.

The results showed a significant positive link between the NASDAQ index with the USA's GDP, the consumer opinion survey, and the consumer price index, which means that an increase of one of these factors increased the tech market index. The NASDAQ also had a positive but less significant correlation with the unemployment rate. Regarding the other macroeconomic factors, the market value of marketable treasury debt was significantly negatively correlated with the NASDAQ index, while the federal funds effective rate and the consumer sentiment index displayed a negative but not statistically significant correlation.

In addition, the model got an adjusted R-squared value of 0.8648, which means that the model fitted the data well and that a large portion of the variability of the NASDAQ composite index was explained. Because this value is unusually high, we performed a model diagnostic, and the results showed some concerns about the reliability of the model. These concerns include the presence of influential outliers and non-normality of the residuals. This should lead to further examination of these findings in the future.

To complete this analysis, we made a correlation matrix to visualize the interplay between these variables (Figure 15).

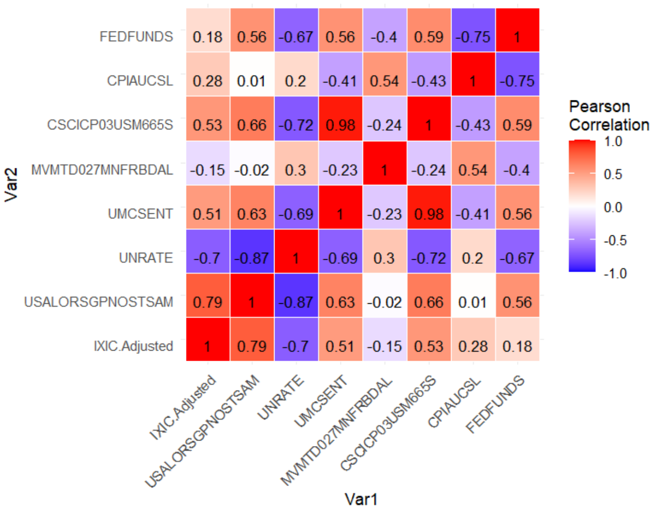

Figure 15: Correlation Matrix between NASDAQ and Macroeconomic Factors, 1995-2005

We observe that the correlations go in the same sense as in the linear model used before, except for three indicators (federal funds rate, consumer sentiment, and unemployment rate) which were not statistically significant.

In conclusion, the NASDAQ index had a positive correlation with the USA's GDP, the consumer opinion survey, and the consumer price index. This confirms that the technology market evolves positively with the overall economic growth and individual's confidence, which is in line with the literature review which emphasizes the influence of behavioral finance on stock bubbles. In addition, the marketable treasury debt was negatively correlated with the tech market, potentially indicating that investors preferred safest government securities in periods of crisis.

4. Impact of Stock Bubbles on Market Stability

4.1 Volatility Analysis

During episodes of bubbles, volatility is often high around this period. Market volatility refers to the ways stock prices change within a certain timeframe.

4.1.1 Implications of Volatility and Broader Market Impact

There are multiple implications linked to volatility. These include an increase in the risk perception of investors, which might want a higher risk premium to compensate for it. Linked to that first point, the behavior of investors is also influenced and could amplify this volatility because of their fear. They could also lose confidence and go from over-optimism to over-pessimism. In addition, Jacobsen & Dannenburg (2003) highlight in their paper the phenomena of volatility clustering and the fact that a high volatility can create similar episodes shortly after. Their paper also highlights the use of the GARCH model, which will be explored in the following sections. Volatility can also have broader implications, impacting multiple markets. A bubble and a high volatility in an asset can contaminate other assets from the same sector, and even other sectors unrelated. This contagion can reduce the benefits of diversification in investment. This will be analyzed in a further section.

4.1.2 GARCH Model and Correlation Analysis

The Generalized Autoregressive Conditional Heteroskedasticity model is commonly used to assess the level of volatility and volatility clustering, as related in the paper of Jacobsen & Dannenburg (2003), by comparing long-term volatility and short-term fluctuations. This model is an extension of the ARCH model of Engle (1982) by Bollerslev (1986). We still focus on the period around the dot-com bubble, between 1995 and 2005, with a focus on MSFT, the S&P 500 and the NASDAQ. Section a) focuses on day-to-day changes, while Section b) focuses on daily returns, which are more common in financial analysis. The model can be represented as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

where σ_t^2 is the variance, r_{t-1}^2 is the lagged squared return, and α_0, α_1, β_1, are parameters to estimate (Bollerslev, 1986).

a) Day-to-Day Changes

The computed values of these parameters can be found in Appendix 7. One notable thing is that β_1 is relatively high and $\alpha + \beta$ is close to 1 which indicates that past volatility influenced future volatilities and that significant changes in volatility had a lasting effect.

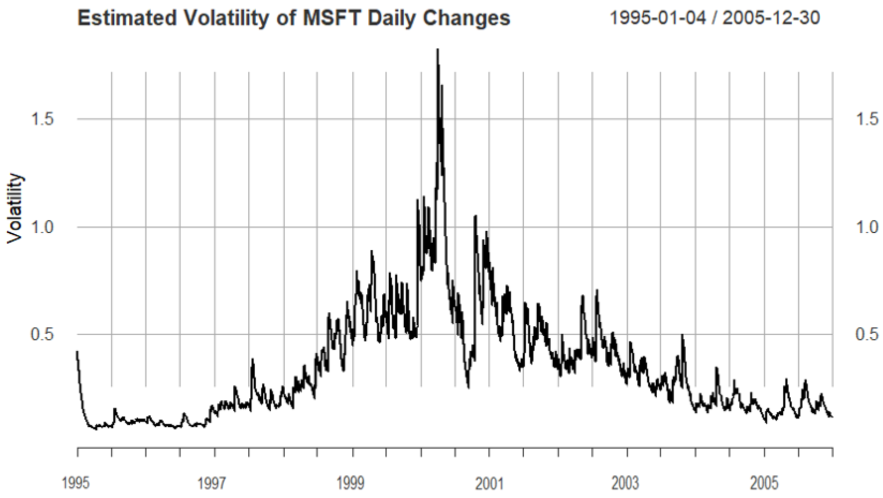

Figure 16: MSFT Daily Changes/Volatility, 1995-2005

Figure 16 further confirmed our finding from the literature. Indeed, we can observe a high volatility around the bubble and that this volatility increases and decreases gradually, proving long-term effect. We observe a volatility between 5.74% on the quietest day, and 182.62% on the most volatile day. Half of the volatility values in this period fall between approximately 15.5% and 48.31%. On average, the estimated volatility was about 34%.

We performed the same analysis for the S&P 500 to see if there are some patterns (Figure 17). The same peak appears in the two graphs, around 2000, providing further evidence of the impact of a bubble on the stability of the whole market.

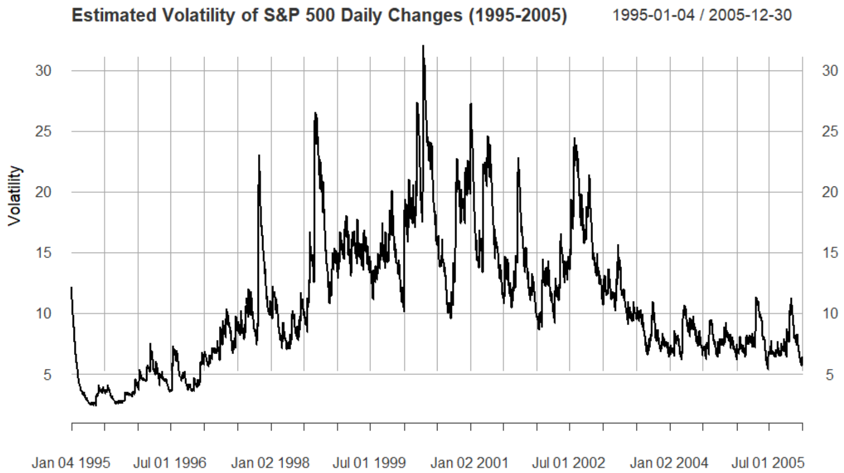

Figure 17: S&P 500 Daily Changes/Volatility, 1995-2005

However, the analysis of day-to-day change is for contextual purpose and graph analysis and is not the most relevant. Indeed, we can observe a significant difference in the values of volatility between MSFT and the S&P 500, due to the difference of their price at that moment. This is not the case for daily return analysis, which is more relevant.

b) Daily Return

The same analysis was performed on daily returns with MSFT, S&P 500 and NASDAQ (Figure 18).

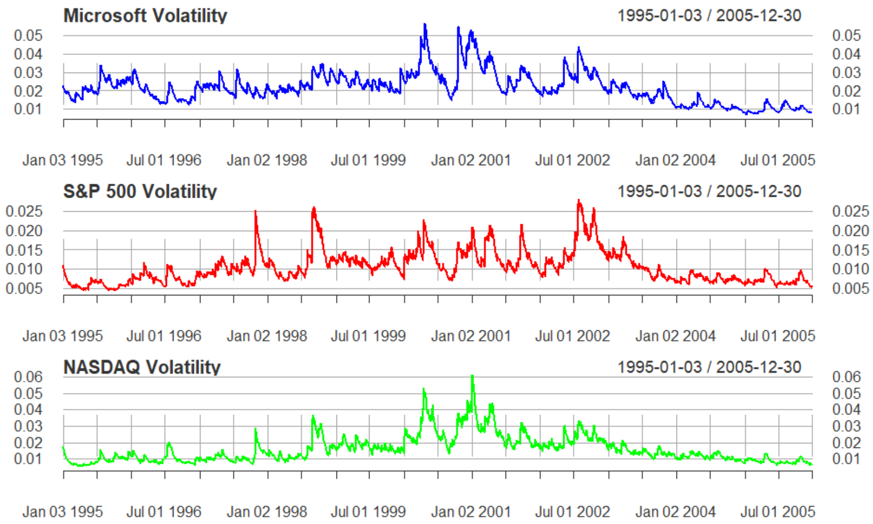

Figure 18: Comparison Volatility of MSFT, S&P 500 & NASDAQ, 1995-2005

The correlations were 0.6730546 for Microsoft and the S&P 500, 0.8083005 for Microsoft and NASDAQ, and 0.7994434 for the S&P 500 and the NASDAQ. These results show that there existed a strong positive correlation between the S&P 500 and the Nasdaq during the period of study. This further confirmed the impact of the stock bubble, thus increasing volatility on the whole market and its stability. However, it is observable that the values of volatility are higher for Microsoft and the NASDAQ than for the whole S&P 500, due to the diversity of sectors in the index.

4.2 Propagation of Stock Bubbles and Interdependence between Asset Classes

It is known that asset classes are not perfectly independent from each other. Even if holding different assets could be a way to diversify a portfolio, their interdependence can be exacerbated in case of crisis. We already analyzed the correlation of volatility between market indexes such as the NASDAQ and S&P 500 in the previous section and observed a strong correlation between them. Still considering the dot-com bubble as a case in point, we made a correlation matrix using R between the NASDAQ, treasury bonds, gold, international assets, and real estate (Figure 19). We used indexes as a proxy, retrieved from Yahoo Finance (2023): ^IXIC, ^TNX, GC=F, IYR, and EFA, from January 1, 1998, to December 31, 2002.

Figure 19: Correlation Matrix of Asset Classes, 1998-2002

We can observe a positive correlation of 0.42 between 10-years treasury bonds and NASDAQ, indicating movement in the same sense for these two asset classes around the dot-com bubble. The correlation between NASDAQ and real estate was also moderately positive, showing the same trend. Regarding the correlation of this tech-heavy index and gold, we can observe a light negative correlation of -0.22. Gold is, in fact, negatively correlated with every index we analyzed around this period, aligning with the theory that highlights the use of gold as a hedge. Finally, the matrix shows a strong correlation with NASDAQ and international markets, providing further evidence of the connection between global markets and the potential repercussions of stock bubbles.

In today’s world, globalization has become increasingly present. This globalization increases the ease for crises to be propagated. Factors influencing it include the interconnection of banks and financial institutions, which can invest in bubble-affected markets abroad, and the organizations that import and export goods, which could be affected by a crisis in their country.

To deepen the analysis, we looked at the correlation between the NASDAQ (^IXIC), a European index (^GDAXI) and a tech-heavy index in Asia (KQ11) between 1998 and 2002. We got a positive correlation between the NASDAQ and Germany’s DAX of 0.54, which further provides evidence of the link between international markets. Regarding South Korea’s KOSDAQ and NASDAQ, we got a low correlation of 0.18. While it doesn't indicate an independence between these markets, the low value could be explained by the fact that we didn’t find data before late 2000 and by some local factors in Asia. We plotted a graph with the standardized closing price of each index, to provide visual evidence of the markets' trends around the dot-com bubble (Figure 20).

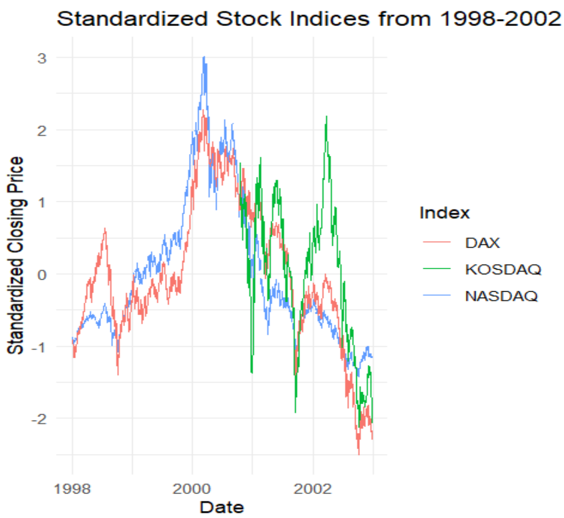

Figure 20: Standardized Closing Price of DAX, KOSDAQ & NASDAQ, 1998-2002

These correlations, even if they don’t mean causality, provide indications regarding the propagation of the crisis linked to the dot-com bubble. We can observe links between these asset classes, especially between international markets. It emphasizes the need to detect these phenomena as early as possible. However, we should keep in mind that these correlations were computed using indexes, which could be biased and not accurately represent the whole market or the asset class they represent.

5. Portfolios' Responses

Even if we understand how to detect stock bubbles and are aware of their potential impacts on market stability, investing prudently remains crucial. While many risk management strategies exist, portfolios could still be significantly impacted. Given the wide scope of this subject, we chose to focus only on individual investments. For consistency with the other sections, this section will principally be based on the historical dot-com bubble.

In this section, we constructed four different types of portfolios to back-test them during the period from 1995 to 2005. The goal of this analysis is to better understand which asset allocation strategies would perform the best and to what extent a huge bubble could impact them. We took portfolios of $10,000 allocation each, supposing no cash reserve and no rebalancing for 10 years. We used adjusted close price, as in most of the other analysis, to account for dividend and other corporate actions. Therefore, we assumed that dividends were reinvested in the stock where they come from, increasing significantly returns compared to withdrawing the dividend in cash. All data were retrieved from Yahoo Finance (2023), except for gold which has been used on a yearly basis and retrieved from the National Mining Association, nma.org (2023). All the computations have been performed on Excel.

5.1 Portfolios Construction

5.1.1 Recall on Diversification

Diversification is one of the most fundamental principles in investment. It means holding different assets with a low or negative correlation between each other. By doing this, if an asset underperforms, the others could stay stable or increase, reducing the potential loss. A well-diversified portfolio can offer a better risk-return trade-off. Diversification could be done via different strategies such as asset class diversification (equities, commodities, real estate, fixed income, …), sectoral diversification (technology, healthcare, energy, …), or geographical diversification.

As argued by Shiller (n.d) in his course "Financial Markets" at Yale University, one of the most notorious examples of hedge is the negative correlation between stocks and gold. To assess the efficiency of this hedge in time of bubble, we plotted the evolution of the price of gold and the adjusted close price of the S&P 500 around the period of the dot-com bubble.

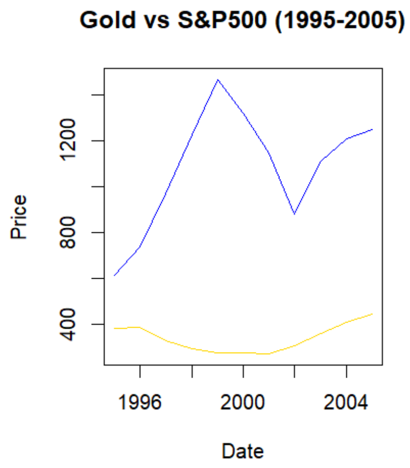

Figure 21: Evolution of Gold Prices vs. S&P 500 Prices, 1995-2005

The blue line in Figure 21 represents the evolution of the S&P 500 prices and the yellow line the gold prices evolution. We computed a correlation of -0.336284 between these two assets in this period. This indicates that they move downward, but not perfectly. This means that gold is probably not a perfect hedge against stocks in times of bubbles but is still useful. However, we keep in mind that this correlation is valuable only around this period and can vary from one time to another. Also, we should note that, as Faber (2007) highlighted in his paper "A Quantitative Approach to Tactical Asset Allocation", the benefits of diversification may be reduced or even disappear in case of a bubble burst (like the one in 2008), as non-correlated assets decline simultaneously.

5.1.2 Portfolio Types

The four portfolio types used are:

- Aggressive growth portfolio, tech-heavy. It is composed of leading tech companies of this period such as Microsoft, Cisco, and Intel. This portfolio's objective is to reach high returns.
- Balanced portfolio. It is composed of different assets including a mix of equities from different sectors, bonds, and other assets. This portfolio's objective is to have a balance between capital appreciation and income generation.
- Conservative portfolio. It is composed of bonds and equities, trying to reach a good income generation while preserving capital, with a minimum risk.

- International diversified portfolio. It is composed of assets across international markets. This portfolio's objective is to reach capital appreciation through international diversification and investment in emerging markets.

5.1.3 Asset Allocation

a) Aggressive Growth Portfolio

This portfolio is heavy on tech and internet stocks. It is composed of:

- Tech Stocks: Microsoft (MSFT) - 20%, Cisco (CSCO) - 20%, Intel (INTC) - 20%, Oracle (ORCL) - 15%, Qualcomm (QCOM) - 10%, Adobe (ADBE) - 5%, Xerox Corporation (XRX) - 5%.
- Non-Tech Stocks: Home Depot (HD) - 5%.

b) Balanced Portfolio

This portfolio is focused on a mix of growth and stability. It is composed of:

- Blue-chip Stocks: Procter & Gamble (PG) - 10%, Johnson & Johnson (JNJ) - 10%, Coca-Cola (KO) - 10%.
- Bonds: Vanguard Total Bond Market Index Fund (VBMFX) - 30%.
- Real Estate: Realty Income Corporation (O) - 10%.
- Tech Stocks: Apple (AAPL) - 10%, IBM (IBM) - 10%.
- Commodities: Gold - 10%.

c) Conservative Portfolio

This portfolio is focused on a mix of stability and income. It is composed of:

- Utilities Stocks: Duke Energy (DUK) - 15%, Southern Company (SO) - 15%.
- Consumer Staples Stocks: Walmart (WMT) - 15%, PepsiCo (PEP) - 15%.
- Bonds: Vanguard Long-Term Treasury Fund (VUSTX) - 15%, PIMCO Total Return Fund (PTTAX) - 15%.
- Dividend-Paying Blue Chips: ExxonMobil (XOM) - 5%.
- Commodities: Gold - 5%.

d) Diversified International Portfolio

This portfolio has a diverse international exposure. It is composed of:

- Emerging Markets: Vanguard Emerging Markets Stock Index Fund (VEIEX) - 21%.
- European Stocks: Unilever (UL) - 14%, AstraZeneca plc (AZN) - 14%.
- Asian Stocks: Toyota Motor Corp (TM) - 12%, Panasonic Corporation (PCRFY) - 12%.

- Bonds: Templeton Global Bond Fund (TPINX) - 14%.
- Commodities: Gold - 8%.
- Tech Stock: Microsoft (MSFT) - 5%.

5.2 Back-testing & Analysis

5.2.1 Back-testing

We computed the number of shares of each asset and the everyday value of each portfolio assuming that the portfolio's holder doesn't touch anything for ten years. Figure 22 represents the evolution of the portfolios during this period:

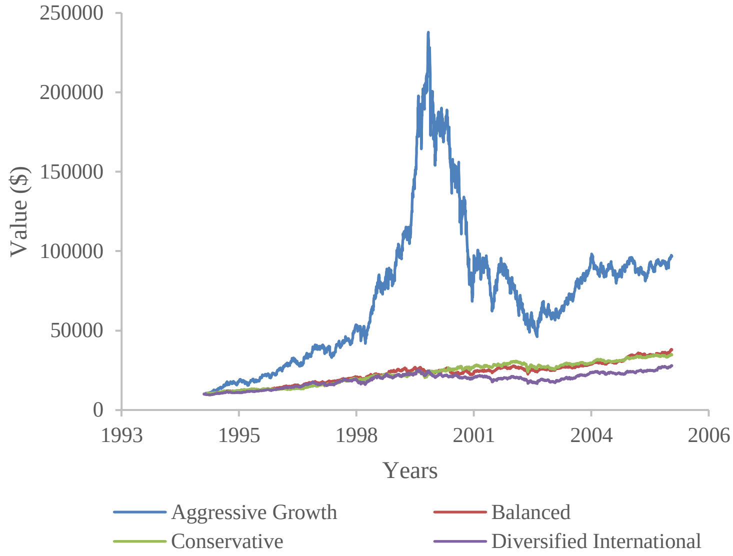

Figure 22: Evolution of 4 Hypothetical Portfolios, 1995-2005

The aggressive portfolio ended with $96,691.10, the balanced one with $37,884.41, the conservative one with $34,612.02 and the diversified international one with $27,734.01. It is observable that the tech-heavy portfolio largely outperformed the others, even if it has experienced a significant decrease during the dot-com bubble. This is logical knowing the long-term perspective, and it confirms the classical idea that equities outperform other assets over long horizons. However, we should note that if investors had invested in this kind of portfolio after the half of 1999, their results would have been null or negative. Even if this portfolio has returned more than the double of the others in this analysis, it is not suitable for a prudent investor, or any risk-averse person.

To better see which portfolio fitted the best to perform well even when an unexpected bad event (bubbles) happened, we looked closer at the others (Figure 23).

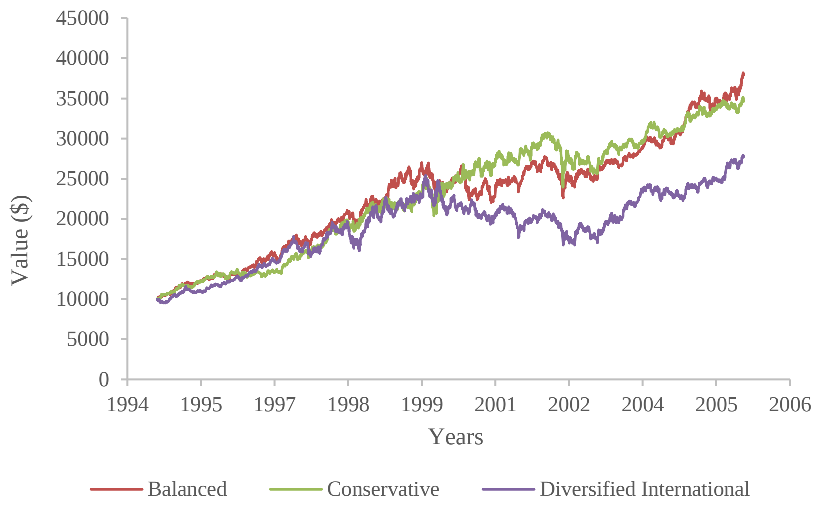

Figure 23: Evolution of 3 Hypothetical Portfolios, 1995-2005

We can notice that the growth of these portfolios was more stable than the previous one, which indicates lower risk and a better tolerance to the dot-com bubble. In addition, it is observable that, even if there have been lower performance periods, an investor could have bought one of these portfolios at any moment and still get a positive return at the end of 2005. However, we can observe that the “diversified international” underperformed the others: its final return is lower, and the growth has been unstable. This could be explained by various international crises during this period, such as the Asian Financial Crisis, which underscored the risks of global investments. It is also consistent with Section 4.2 where we found positive correlations between international markets around the dot-com bubble. The two best performing portfolios, which have been performing well most of the time and have not been impacted significantly by the bubble burst, are the balanced one and the conservative one.

5.2.2 Comparison with a Benchmark: S&P 500

The purpose of this section is to know if a portfolio, well-balanced or conservative, performs better than simply buying an S&P 500 ETF and holding it forever, as some investors suggest. The following graph represents their evolutions, with still the same initial investment of $10,000. For this analysis, we took the SPDR S&P 500 ETF Trust (SPY) index.

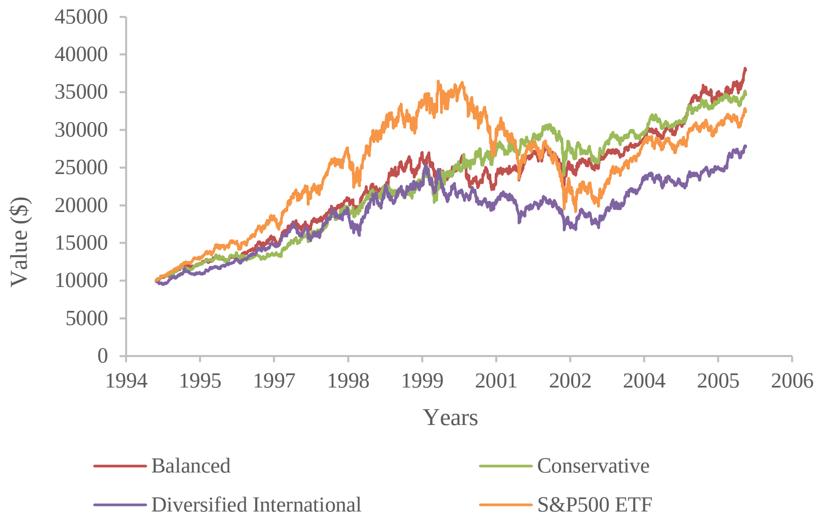

Figure 24: Evolution of 3 Hypothetical Portfolios and an S&P 500 Index, 1995-2005

It is observable that the S&P 500 ETF, at the end of the run, had underperformed the two best portfolios, with a final return of $32,393.59 (Figure 24). In addition, this ETF, as well as the aggressive growth portfolio, had been negatively impacted by the dot-com bubble. This is logical knowing that the S&P 500 contains a lot of tech companies. However, unlike the tech-heavy portfolio, its final return isn't worth it compared to the other portfolios. However, we keep in mind that the results could be different in a period without a crisis.

5.2.3 Comparison with a Risk-Free Investment / Sharpe Ratio

The Sharpe ratio is a tool used to provide information about the risk-adjusted performance of an investment or portfolio. It helps understand how well a portfolio's returns compensate for the level of risk taken. We used this metric to assess the performance of the balanced portfolio, the best-performing in the backtesting analysis, against a risk-free rate. We chose this portfolio rather than the conservative one, which also performs well, because of the level of risk a bit higher in the "balanced portfolio". The formula employed for this purpose was:

$$Sharpe\ Ratio = \frac{Portfolio\ Return - Risk\ Free\ Rate}{Standard\ Deviation\ of\ portfolio's\ excess\ return}.$$

We chose to take a U.S. Treasury Security of 10 years, as it is the same time period as the period we were looking at with the portfolios. The data of this Risk-Free Rate were retrieved from the Economic Report of the President, as documented by the Council of Economic Advisers, Executive Office of the President (2011).

Overall, we got a positive Sharpe ratio for every year, except for 2000 and 2002. This can be explained by the burst of the dot-com bubble. Given the balanced portfolio's exposure to tech stocks such as Apple and IBM, this event has adversely impacted the portfolio's performance. In addition, another factor explaining these negative values is the fact that the risk-free rate was higher in 2000. Finally, the high volatility in these years, especially in 2000 when the standard deviation was the highest of the period, could have negatively influenced the Sharpe ratio. To conclude, while the balance portfolio showed strong performance during most of the years under study, it indicated some vulnerabilities in case of important events such as the dot-com bubble. A table with the year-by-year computation of Sharpe ratios could be found in Appendix 11.

5.2.4 Conclusion and Limitations

In conclusion, this analysis has provided evidence of what was found in the literature: that diversification is probably the most fundamental risk mitigation strategy. It also proves that we can get a positive return with these types of portfolios, even in periods of crisis and high volatility. In addition, we should note that long term perspective is key in investing, especially in this kind of period. However, we should acknowledge some potential limitations due to bias. Indeed, as the period of 1995-2005 is far from now, even if we tried to retrieve notorious organizations for that period, we could have chosen those that we know the most in the present. In addition, we used ETFs for this analysis to represent sectors like bonds or real estate, and ETFs don't perfectly represent the assets they are tracking.

6. Conclusion

To conclude this master's thesis, we gained insight into the phenomena of stock bubbles and their implications for the market stability and investment portfolios. For this purpose, the application of bubble detection algorithms such as the ADF, PP, and KPSS tests was used to test the stationarity of the time series. We found that the data tested with Microsoft were not stationary, which is usual for this type of financial series. We then applied a GSADF test, following the method of Phillips, Shi, and Yu (2015), a structural break test, and a BDS test. Combined with these methodologies, we searched for deviations from historical trends and fundamental valuations to deepen the analysis. For this purpose, we compared historical stock prices with a 90-day rolling average, we looked at P/E ratio, dividend yield and returns distributions. In addition to these quantitative methodologies to detect stock bubbles that we applied around the period of the dot-com bubble, we also deepened our understanding of the factors contributing to these phenomena through more literature research. We then also applied a test on R to evaluate the duration and intensity of the dot-com bubble on Microsoft. We also used a regression to see which macroeconomic indicators have influenced the forming of the bubble. We found a positive correlation with the USA's GDP, the consumer opinion survey, and the consumer price index, in line with the literature review which highlights the influence of behavioral finance in the forming stock bubbles.

This thesis continues with research on the impact of these speculative phenomena on market stability, where we gained insight into the volatility around the dot-com bubble using GARCH models on day-to-day change and daily returns. We found that the volatilities of Microsoft, the S&P 500, and the NASDAQ were highly correlated around the period of the dot-com bubble, highlighting the propagation of the tech bubble to the broader markets. To deepen the analysis of the propagation, we made a correlation matrix between different asset classes between 1998 and 2002, and found positive correlations between every asset class (bond, tech stock market, real estate, international stock market) except for gold. We also performed a short analysis on tech markets over different countries and found a positive correlation. This highlights the importance of policymakers and regulators, as well as the need to detect these phenomena as early as possible, to prevent and mitigate the adverse effects.

We closed this thesis by looking at investment portfolios. After a recall on diversification, we constructed four different portfolio types: an aggressive growth one, a balanced one, a conservative one, and a diversified international one. These portfolios were back-tested between 1995 and 2005 to see which one performed the best around the period of the dot-com bubble, and then compared with a benchmark (the S&P 500). We found that the two best performing portfolios were the balanced one and the conservative one. We tested the balanced portfolio against a risk-free rate, by using the Sharpe ratio, and found that the risk-return tradeoff was worth it 9 years out of 11. This highlights the benefits of a well-balanced portfolio rather than a risk-free investment in the long run, even when a crisis happens.

In conclusion, this thesis introduces guidelines regarding the ways to mitigate stock bubbles, by learning how to analyze patterns to detect them, how to quantify the impact on broader markets' stability, and how to make resilient investment portfolios. By looking at both theoretical frameworks and empirical analysis, we highlighted the importance of employing a diverse panel of tools and economic indicators, from quantitative analysis to qualitative analysis.

7. Bibliography

Adobe Inc. (2023). Investor relations: Financial reports. Retrieved May 23, 2023, from https://www.adobe.com/investor-relations.html.

Ali, A., Jan, F. A., & Jehanzeb, M. (2015). Effects of stock market crash on investors behavior. *Journal of management info, 2*(2), 13-21. https://doi.org/10.31580/jmi.v6i1.42.

Bikhchandani, S., & Sharma, S. (2001). Herd behavior in financial markets. *IMF Staff papers, 47*(3), 279-310.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics, 31*(3), 307-327. https://doi.org/10.1016/0304-4076(86)90063-1.

Board of Governors of the Federal Reserve System (US). (2024). *Federal funds effective rate* [FEDFUNDS]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/FEDFUNDS.

Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. *Journal of Econometrics, 108*(2), 343-363. https://doi.org/10.1016/S0304-4076(01)00139-7.

Brock, W., Dechert, D., Scheinkman, J., & LeBaron, B. (1996). A Test for Independence Based on the Correlation Dimension. *Econometric Reviews, 15*(3), 197-235. https://doi.org/10.1080/07474939608800353.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological), 37*(2), 149-192. https://doi.org/10.1111/j.2517-6161.1975.tb01532.x.

Council of Economic Advisers, Executive Office of the President. (2011, February 23). *Econ. Rept. 2011 - Table B-73: Bond yields and interest rates, 1933-2010*. U.S. Government Printing Office. https://www.govinfo.gov/app/details/ERP-2011/ERP-2011-table73.

Cuñado, J., Gil-Alana, L. A., & De Gracia, F. P. (2009). US stock market volatility persistence: evidence before and after the burst of the IT bubble. *Review of Quantitative Finance and Accounting, 33*(3), 233-252. https://doi.org/10.1007/s11156-009-0111-5.

Damodaran, A. (2012). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*. Wiley.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association, 74*(366a), 427-431. https://doi.org/10.1080/01621459.1979.10482531.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society, 50*(4), 987-1007. https://doi.org/10.2307/1912773.

Faber, M. T. (2007). A Quantitative Approach to Tactical Asset Allocation. *The Journal of Wealth Management,* Spring 2007. Updated February 2009, February 2013. Retrieved from https://urban-stocks.com/images/pdf/AlleAnderen/Meb-Faber_-A-Quantitative-Approach-to-Tactical-Asset-Allocation.pdf.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, *25*(2), 383-417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x.

Federal Reserve Bank of Dallas. (2024). *Market value of marketable treasury debt* [MVMTD027MNFRBDAL]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/MVMTD027MNFRBDAL.

Geanakoplos, J. (2010). The Leverage Cycle. *NBER Macroeconomics Annual 2009, 24*(1), 1-65.

Graham, B. (2003). *The Intelligent Investor: The definitive book on value investing. A book of practical counsel (Revised ed., J. Zweig, Commentator)*. HarperBusiness.

Graham, B., & Dodd, D. (2008). *Security Analysis: Principles and Technique* (6th ed.). McGraw-Hill.

Hansen, B. E. (2001). The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity. *Journal Of Economic Perspectives*, *15*(4), 117-128. https://doi.org/10.1257/jep.15.4.117.

Ibbotson, R. G., & Kaplan, P. D. (2000). Does asset allocation policy explain 40, 90, or 100 percent of performance? *Financial Analysts Journal*, *56*(1), 26-33. https://doi.org/10.2469/faj.v56.n1.2327.

Intel Corporation. (2023). *Annual Reports*. Retrieved September 23, 2023, from https://www.intc.com/filings-reports/annual-reports.

Jacobsen, B., & Dannenburg, D. (2003). Volatility clustering in monthly stock returns. *Journal of Empirical Finance*, *10*(4), 479–503. https://doi.org/10.1016/s0927-5398(02)00071-3.

Kindleberger, C. P., & Aliber, R. Z. (2005). *Manias, Panics, and Crashes: A History of Financial Crises* (5th ed.). John Wiley & Sons, Inc.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1-3), 159-178. https://doi.org/10.1016/0304-4076(92)90104-Y.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77-91. https://doi.org/10.2307/2975974.

Miao, J., & Wang, P. (2014). Sectoral bubbles, misallocation, and endogenous growth. *Journal of Mathematical Economics*, *53*, 153–163. https://doi.org/10.1016/j.jmateco.2013.12.003.

National Mining Association. (n.d.). *Historical Gold Prices - 1833 to Present*. Retrieved November 15, 2023, from https://nma.org/wp-content/uploads/2016/09/historic_gold_prices_1833_pres.pdf.

Okina, K. &. Shirakawa, M. &. Shiratsuka, S. (2001). The Asset Price Bubble and Monetary Policy: Japan's Experience in the Late 1980s and the Lessons: Background paper. *Monetary and Economic Studies*, *19*(S1), 395-450.

Organization for Economic Co-operation and Development. (2023). *Consumer Opinion Surveys: Confidence Indicators: Composite Indicators: OECD Indicator for United States* [CSCICP03USM665S]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/CSCICP03USM665S.

Organization for Economic Co-operation and Development. (2024). *Leading indicators OECD: Reference series: Gross Domestic Product (GDP): Normalised for United States* [USALORSGPNOSTSAM]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/USALORSGPNOSTSAM.

Page, E.S. (1954). Continuous Inspection Schemes. *Biometrika*, *41*(1/2), 100-115. https://doi.org/10.2307/2333009.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-346. https://doi.org/10.1093/biomet/75.2.335.

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. https://doi.org/10.1111/j.1468-2354.2010.00625.x.

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. https://doi.org/10.1111/iere.12132.

Schwartz, A.J. (2009). Origins of the Financial Market Crisis of 2008. Cato Journal, 29(1), 19-23.

Shiller, R. (n.d.). *Financial Markets*. Coursera. Accessed 2023. https://www.coursera.org/learn/financial-markets-global.

Shiller, R. J. (2000). *Irrational Exuberance*. Princeton University Press.

Shiller, R. J. (2002). Bubbles, Human Judgment, and Expert Opinion. *Financial Analysts Journal*, *58*(3), 18-26. https://doi.org/10.2469/faj.v58.n3.2535.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, *5*, 297-323. https://doi.org/10.1007/bf00122574.

University of Michigan. (2024). *University of Michigan: Consumer sentiment* [UMCSENT]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/UMCSENT.

UpMyInterest (2023), *Summary*. Retrieved November 17, 2023, from https://www.upmyinterest.com/sp500-pe/.

U.S. Bureau of Labor Statistics. (2024). *Consumer Price Index for All Urban Consumers: All Items in U.S. City Average* [CPIAUCSL]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/CPIAUCSL.

U.S. Bureau of Labor Statistics. (2024). *Unemployment rate* [UNRATE]. Federal Reserve Bank of St. Louis. Retrieved March 28, 2024, from https://fred.stlouisfed.org/series/UNRATE.

Yahoo Finance. (n.d.). Retrieved from https://finance.yahoo.com.

YCharts (2023). *Microsoft Corp (MSFT) EPS Diluted (Annual)*. Retrieved November 14, 2023, from https://ycharts.com/companies/MSFT/eps_annual.

8. Appendices

Appendix 1: Bubble Detection algorithm: R Script (3.1.1)

a) Data Collection

```
library(quantmod)

getSymbols("MSFT", src = "yahoo", from = "1995-01-01", to = "2005-12-31")

msft_price <- Cl(MSFT)
```

b) Stationarity Tests

ba) ADF test

```
library(urca)

adf_test <- ur.df(msft_price, type="drift")

summary(adf_test)
```

Result:

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

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression drift

Call:

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

Residuals:

    Min      1Q  Median      3Q     Max

-7.6023 -0.2346 -0.0240  0.2312  5.0708

Coefficients:

             Estimate Std. Error t value
```

```
(Intercept)  0.062873   0.029980   2.097

z.lag.1     -0.002211   0.001096  -2.017

z.diff.lag  -0.021319   0.018999  -1.122

            Pr(>|t|)

(Intercept)   0.0361 *

z.lag.1       0.0438 *

z.diff.lag    0.2619

---

Signif. codes:

  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

  0.1 ' ' 1

Residual standard error: 0.6772 on 2766 degrees of freedom

Multiple R-squared:  0.001954,      Adjusted R-squared:  0.001233

F-statistic: 2.708 on 2 and 2766 DF,  p-value: 0.06685

Value of test-statistic is: -2.0172 2.2403

Critical values for test statistics:

      1pct  5pct 10pct

tau2 -3.43 -2.86 -2.57

phi1  6.43  4.59  3.78
```

bb) PP Test

```
pp_test <- ur.pp(msft_price, type="Z-tau")

summary(pp_test)
```

bc) KPSS Test

```
library(urca)
```

```
kpss_test <- ur.kpss(data$MSFT.Adjusted, type="tau")

summary(kpss_test)
```

Result:

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: tau with 9 lags.

Value of test-statistic is: 5.0697

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

bd) ADF Test on Differenced Data

```
df_test <- ur.df(diff(msft_price)[-1], type="drift")

summary(df_test)
```

Result:

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

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression drift

Call:

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



Residuals:

    Min       1Q  Median       3Q      Max
```

```
-7.6137 -0.2361 -0.0090  0.2295  5.0677

Coefficients:

             Estimate Std. Error t value

(Intercept)  0.008057   0.012881   0.625

z.lag.1     -0.994713   0.027180 -36.597

z.diff.lag  -0.026740   0.019011  -1.407

            Pr(>|t|)

(Intercept)    0.532

z.lag.1       <2e-16 ***

z.diff.lag     0.160

---

Signif. codes:

  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

  0.1 ' ' 1

Residual standard error: 0.6776 on 2765 degrees of freedom

Multiple R-squared:  0.5114,  Adjusted R-squared:  0.511

F-statistic:  1447 on 2 and 2765 DF,  p-value: < 2.2e-16

Value of test-statistic is: -36.5969 669.6678

Critical values for test statistics:

      1pct  5pct 10pct

tau2 -3.43 -2.86 -2.57

phi1  6.43  4.59  3.78
```

Graph script:

```
msft_diff <- diff(msft_price) # calculate the differences between
consecutive observations

msft_diff_clean <- na.omit(msft_diff) # remove “NA” values

par(mfrow=c(2,1)) # make the two plots look vertically stacked


# Plot the original price

plot(msft_price, type="l", col="blue", main="Microsoft Stock Price (1995-
2005)", ylab="Price", xlab="Date")


# Plot the difference in price

plot(msft_diff_clean, type="h", col="red", main="Difference in Microsoft
Stock Price (1995-2005)", ylab="Difference", xlab="Date")
```

c) GSADF

```
# Load necessary libraries

library(quantmod)

library(tseries)


# Retrieve Microsoft stock data from Yahoo Finance

getSymbols("MSFT", src = "yahoo", from = "1995-01-01", to = "2005-12-31")


# Extract closing prices and log-transform them

msft_prices <- Cl(MSFT)

log_msft_prices <- log(msft_prices)
```

```
# Define the minimum window size for the GSADF test

min_window_size <- 30

length_data <- length(log_msft_prices)

gsadf_stats <- rep(NA, length_data) # Initialize the vector with NAs to
match the length


# Conduct the GSADF test using an expanding window approach

for (i in seq(min_window_size, length_data)) {

      subset_data <- log_msft_prices[(i-min_window_size+1):i]

      if (sd(subset_data) > 0) {  # Avoiding constant series

      test_result <- adf.test(as.numeric(subset_data), alternative =
"stationary")

      gsadf_stats[i] <- test_result$statistic  # Store the statistic at the
correct index

      }

}


# Clean NA values and determine the threshold for significant GSADF
statistics

clean_gsadf_stats <- gsadf_stats[!is.na(gsadf_stats) & gsadf_stats > 0]

threshold <- quantile(clean_gsadf_stats, 0.95)


# Identify potential bubbles where the GSADF statistic exceeds the
threshold

results <- data.frame(

      Date = index(msft_prices)[(min_window_size+1):length(msft_prices)],
```

```
      GSADF_Stats = gsadf_stats[(min_window_size+1):length(msft_prices)]
)


# Plot the GSADF statistics with potential bubble periods highlighted

plot(results$Date, results$GSADF_Stats, type = "l",

      main = "GSADF Test Statistics for MSFT with Potential Bubbles")

abline(h = threshold, col = "red", lwd = 2) # Horizontal line for the
threshold

potential_bubbles <- results[results$GSADF_Stats > threshold, ]

points(potential_bubbles$Date, potential_bubbles$GSADF_Stats, col = "blue",
pch = 20) # Points for potential bubbles


# Print the dates and GSADF statistics for potential bubbles

print(potential_bubbles)
```

Result:

```
Date GSADF_Stats
724 1997-12-23    1.529061
725 1997-12-24    1.841031
849 1998-06-24    2.006305
```

d) Structural Break Test on Raw Data

```
library(strucchange)

breaks <- breakpoints(msft_price ~ 1)

plot(breaks)
```

Result:

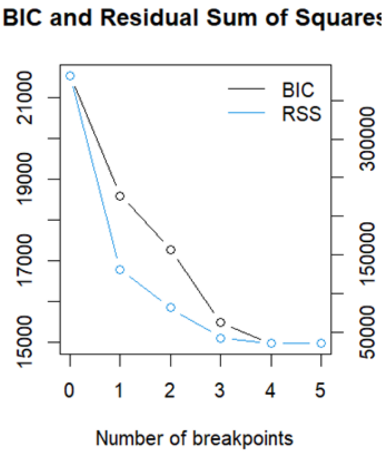

Figure 25: Breakpoints MSFT, 1995-2005

e) Breakpoint Identification

```
total_observations <- length(msft_price)

breakpoint_indices <- round(c(0.207, 0.357, 0.507, 0.658) *
total_observations)

breakpoint_dates <- index(msft_price)[breakpoint_indices]

print(breakpoint_dates)
```

f) BDS Test

```
# Load necessary libraries and retrieve data

library(quantmod)

start_date <- as.Date("1995-01-01")

end_date <- as.Date("2005-12-31")

msft_data <- getSymbols("MSFT", src = "yahoo", from = start_date, to =
end_date, auto.assign = FALSE)


# Apply the BDS test

embedding_dims <- c(2,3,4)
```

```
results <- list()

for (m in embedding_dims) {

     results[[paste0("m=", m)]] <- bds.test(msft_data$MSFT.Adjusted, m = m)

}

results
```

Result:

```
$`m=2`

       BDS Test

data:  msft_data$MSFT.Adjusted

Embedding dimension =  2

Epsilon for close points =   3.6868  7.3735 11.0603 14.7471

Standard Normal =
      [ 3.6868 ] [ 7.3735 ] [ 11.0603 ] [ 14.7471 ]
[ 2 ]   199.9495   179.5715    150.7418    166.6462

p-value =
      [ 3.6868 ] [ 7.3735 ] [ 11.0603 ] [ 14.7471 ]
[ 2 ]          0          0           0           0



$`m=3`

       BDS Test

data:  msft_data$MSFT.Adjusted

Embedding dimension =  2 3

Epsilon for close points =   3.6868  7.3735 11.0603 14.7471
```

```
Standard Normal =
      [ 3.6868 ] [ 7.3735 ] [ 11.0603 ] [ 14.7471 ]
[ 2 ]   200.1299   179.6538    150.7839    166.6310
[ 3 ]   345.7777   230.9509    167.2008    162.4976

p-value =
      [ 3.6868 ] [ 7.3735 ] [ 11.0603 ] [ 14.7471 ]
[ 2 ]          0          0           0           0
[ 3 ]          0          0           0           0



$`m=4`

       BDS Test

data:  msft_data$MSFT.Adjusted

Embedding dimension =  2 3 4

Epsilon for close points =   3.6868  7.3735 11.0603 14.7471

Standard Normal =
      [ 3.6868 ] [ 7.3735 ] [ 11.0603 ] [ 14.7471 ]
[ 2 ]   200.3055   179.7350    150.8240    166.6155
[ 3 ]   346.1899   231.0800    167.2568    162.4850
[ 4 ]   646.5805   306.7346    187.8495    158.7620

p-value =
      [ 3.6868 ] [ 7.3735 ] [ 11.0603 ] [ 14.7471 ]
[ 2 ]          0          0           0           0
[ 3 ]          0          0           0           0
[ 4 ]          0          0           0           0
```

Appendix 2: Deviations from Historical Trends: R Script (3.1.2)

```
# Calculate 90-day rolling average

roll_avg <- zoo::rollmean(msft_price, k=90, align='right', fill=NA)
```

```
# Visualize the stock price alongside the rolling average

# Determine finite limits for ylim

finite_prices <- msft_price[!is.na(msft_price) & !is.infinite(msft_price)]

finite_avg <- roll_avg[!is.na(roll_avg) & !is.infinite(roll_avg)]

y_limits <- c(min(c(finite_prices, finite_avg)), max(c(finite_prices,
finite_avg)))

par(mfrow=c(1,1)) # Reset the plotting area

plot(index(msft_price), msft_price, type="l", col="blue", main="Microsoft
Stock Price (1995-2005) with 90-Day Rolling Average", ylab="Price",
xlab="Date", ylim=y_limits)

lines(index(msft_price), roll_avg, col="red")


# Calculate the difference between the stock price and its 90-day rolling
average

residuals <- msft_price - roll_avg


# Plot the residuals

plot(index(msft_price), residuals, type="h", main="Residuals between MSFT
Price and its 90-Day Rolling Average", ylab="Residuals", xlab="Date")


# Horizontal line at zero for reference

abline(h=0, col="grey")


# Plot threshold of Standard Deviation

mean_residuals <- mean(residuals, na.rm = TRUE)

sd_residuals <- sd(residuals, na.rm = TRUE)
```

```
large_deviation_upper <- mean_residuals + 2*sd_residuals

large_deviation_lower <- mean_residuals - 2*sd_residuals

abline(h = large_deviation_upper, col = "red", lty = 2)

abline(h = large_deviation_lower, col = "red", lty = 2)
```

Appendix 3: Evolution of the P/E ratio of the S&P 500 between 1995 and 2005 (3.1.3)

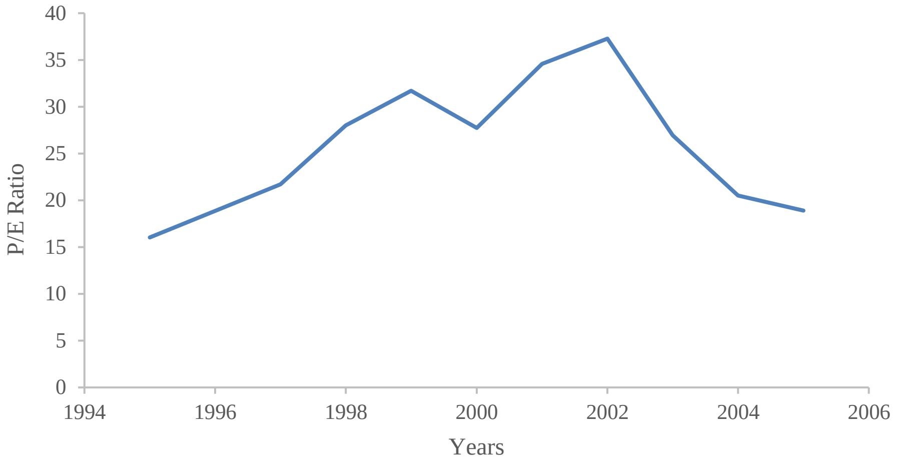

Figure 26: P/E ratio S&P 500, 1995-2005

Appendix 4: Comparative Analysis of Log Returns Distributions: R Script (3.1.4)

Note that these scripts are for MSFT, but the same applies for INTC and ^IXIC.

a) Base script

```
# Load necessary library and set the time period (dot-com bubble)

library(quantmod)

start_date <- "1995-01-01"

end_date <- "2005-12-31"


# Retrieve stock data for Microsoft

getSymbols("MSFT", from = start_date, to = end_date)
```

```
# Extract the adjusted closing prices

stock_data <- Ad(get("MSFT"))


# Calculate log returns

log_returns <- diff(log(stock_data))


# Load necessary library for statistical analysis

library(moments)


# Basic statistical analysis of log returns

mean_return <- mean(log_returns, na.rm = TRUE)

std_deviation <- sd(log_returns, na.rm = TRUE)

skewness <- skewness(log_returns, na.rm = TRUE)

kurtosis <- kurtosis(log_returns, na.rm = TRUE)


# Print results

cat("Mean: ", mean_return, "\nStandard Deviation: ", std_deviation,
"\nSkewness: ", skewness, "\nKurtosis: ", kurtosis, "\n")


# Plot histogram of log returns

hist(log_returns, breaks = 50, main = "Histogram of Log Returns for MSFT",
xlab = "Log Returns", col = "blue")


# Add normal distribution curve
```

```
curve(dnorm(x, mean = mean_return, sd = std_deviation), col = "red", lwd = 2, add = TRUE)

legend("topright", legend = c("Log Returns", "Normal Distribution"), col = c("blue", "red"), lty = 1)


# Retrieve comparative data for a different period

getSymbols("MSFT", from = "2006-01-01", to = "2016-12-31")

comparative_data <- Ad(get("MSFT"))


# Calculate log returns for the comparative data

comparative_log_returns <- diff(log(comparative_data))


# Plot histogram for the comparative period

hist(comparative_log_returns, breaks = 50, main = "Histogram of Log Returns for MSFT (2006-2016)", xlab = "Log Returns", col = "green")


# Add normal distribution curve for the comparative period

comparative_mean <- mean(comparative_log_returns, na.rm = TRUE)

comparative_sd <- sd(comparative_log_returns, na.rm = TRUE)

curve(dnorm(x, mean = comparative_mean, sd = comparative_sd), col = "darkgreen", lwd = 2, add = TRUE)

legend("topright", legend = c("Comparative Log Returns", "Normal Distribution"), col = c("green", "darkgreen"), lty = 1)
```

b) Script for statistical analysis

```
# Calculate and compare the statistical measures

cat("Dot-Com Bubble Period:\n")
```

```
cat("Mean: ", mean(log_returns, na.rm = TRUE), "\n")

cat("Standard Deviation: ", sd(log_returns, na.rm = TRUE), "\n")

cat("Skewness: ", skewness(log_returns, na.rm = TRUE), "\n")

cat("Kurtosis: ", kurtosis(log_returns, na.rm = TRUE), "\n\n")


cat("Post-Bubble Period:\n")

cat("Mean: ", mean(comparative_log_returns, na.rm = TRUE), "\n")

cat("Standard Deviation: ", sd(comparative_log_returns, na.rm = TRUE),
"\n")

cat("Skewness: ", skewness(comparative_log_returns, na.rm = TRUE), "\n")

cat("Kurtosis: ", kurtosis(comparative_log_returns, na.rm = TRUE), "\n")


# Kolmogorov-Smirnov Test

ks_test <- ks.test(log_returns, comparative_log_returns)


# Print KS Test result

print(ks_test)
```

c) Script for another plot

```
# Load necessary library

library(ggplot2)


# Create a combined data frame for plotting

data_for_plot <- data.frame(

  Returns = c(as.numeric(log_returns),
as.numeric(comparative_log_returns)),
```

```
  Period = c(rep("Dot-Com Bubble", length(log_returns)), rep("Post-Bubble", length(comparative_log_returns)))

)


# Remove NA values for better plotting

data_for_plot <- na.omit(data_for_plot)


# Plot density lines for each period

ggplot(data_for_plot, aes(x = Returns, color = Period)) +

  geom_density() +

  scale_color_manual(values = c("Dot-Com Bubble" = "blue", "Post-Bubble" = "green")) +

  labs(title = "Comparative Density Plot of Log Returns",

      x = "Log Returns",

      y = "Density") +

  theme_minimal() +

  theme(legend.title = element_blank())
```

Appendix 5: Duration and Intensity of Bubbles: R Script (3.2.1)

```
# Note that 'residuals' are the differences between stock price and its rolling average

# Detect start of bubble

bubble_starts <- which(residuals > large_deviation_upper)

# Ensure that there are values satisfying the condition

if (length(bubble_starts) > 0) {

  bubble_start <- bubble_starts[1]
```

```
  # Detect end (burst) of bubble after the start

  bubble_ends <- which(residuals[bubble_start:length(residuals)] <
large_deviation_lower)

  if (length(bubble_ends) > 0) {

    bubble_end <- bubble_ends[1] + bubble_start - 1

    bubble_duration <- bubble_end - bubble_start

    print(paste("Duration of the bubble:", bubble_duration, "days"))

    peak_during_bubble <- max(msft_price[bubble_start:bubble_end])

    start_price <- msft_price[bubble_start]

    intensity <- ((peak_during_bubble - start_price) / start_price) * 100

    print(paste("Intensity of the bubble:", round(intensity, 2), "%"))

  } else {

    print("The bubble did not burst within the data range.")

  }

} else {

  print("No bubble detected within the data range.")

}
```

Appendix 6: LM Model NASDAQ and Macroeconomic Indicators: R Script (3.2.3)

a) LM Model

```
# Install and load necessary packages

packages_needed <- c("quantmod", "zoo", "lubridate", "tseries", "car",
"MASS", "lmtest", "sandwich", "caret", "readxl")

packages_to_install <- packages_needed[!(packages_needed %in%
installed.packages()[,"Package"])]
```

```
if(length(packages_to_install)) install.packages(packages_to_install)
lapply(packages_needed, library, character.only = TRUE)

# Retrieve NASDAQ Composite Index data from Yahoo Finance
getSymbols("^IXIC", src = "yahoo", from = "1995-01-01", to = "2005-12-31")
index_data <- Ad(IXIC)

# Convert to monthly data
index_monthly <- to.monthly(index_data, indexAt = "lastof", OHLC = FALSE)

# Import and prepare macroeconomic data
macro_data <- read_excel("C:/Users/32492/Dropbox/Papier étudiant/master's
thesis/data pr modele avec macro factors/data_merge.xlsx")
macro_data$Date <- as.Date(macro_data[[1]])
index_monthly_df <- as.data.frame(index_monthly)
index_monthly_df$Date <- as.Date(rownames(index_monthly_df))
rownames(index_monthly_df) <- NULL
combined_data <- merge(index_monthly_df, macro_data, by = "Date", all =
TRUE)

# Clean NA values
combined_data_clean <- na.locf(combined_data)

# Linear regression model
```

```
lm_model <- lm(IXIC.Adjusted ~ USALORSGPNOSTSAM + UNRATE + UMCSENT +
MVMTD027MNFRBDAL + CSCICP03USM665S + CPIAUCSL + FEDFUNDS, data =
combined_data_clean)

summary(lm_model)


# Model Diagnostics

par(mfrow=c(2,2))

plot(lm_model)
```

Result:

```
Call:

lm(formula = IXIC.Adjusted ~ USALORSGPNOSTSAM + UNRATE + UMCSENT +

    MVMTD027MNFRBDAL + CSCICP03USM665S + CPIAUCSL + FEDFUNDS,

    data = combined_data_clean)


Residuals:

    Min      1Q  Median      3Q     Max

-875.51 -174.41  -25.42  122.69 1248.38


Coefficients:

                   Estimate Std. Error

(Intercept)      -1.068e+05  1.009e+04

USALORSGPNOSTSAM  6.872e+02  6.373e+01

UNRATE            1.532e+02  8.031e+01

UMCSENT          -1.102e+01  1.020e+01

MVMTD027MNFRBDAL -1.017e+00  7.549e-02
```

```
CSCICP03USM665S   3.797e+02  1.040e+02
CPIAUCSL           3.265e+01  4.866e+00
FEDFUNDS         -4.455e+01  3.564e+01
                 t value Pr(>|t|)
(Intercept)      -10.586  < 2e-16 ***
USALORSGPNOSTSAM  10.782  < 2e-16 ***
UNRATE             1.907 0.057588 .
UMCSENT           -1.081 0.280644
MVMTD027MNFRBDAL -13.465  < 2e-16 ***
CSCICP03USM665S    3.651 0.000317 ***
CPIAUCSL           6.709 1.26e-10 ***
FEDFUNDS          -1.250 0.212438
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
  0.1 ' ' 1

Residual standard error: 292.1 on 255 degrees of freedom
Multiple R-squared:  0.8684,  Adjusted R-squared:  0.8648
F-statistic: 240.4 on 7 and 255 DF,  p-value: < 2.2e-16
```

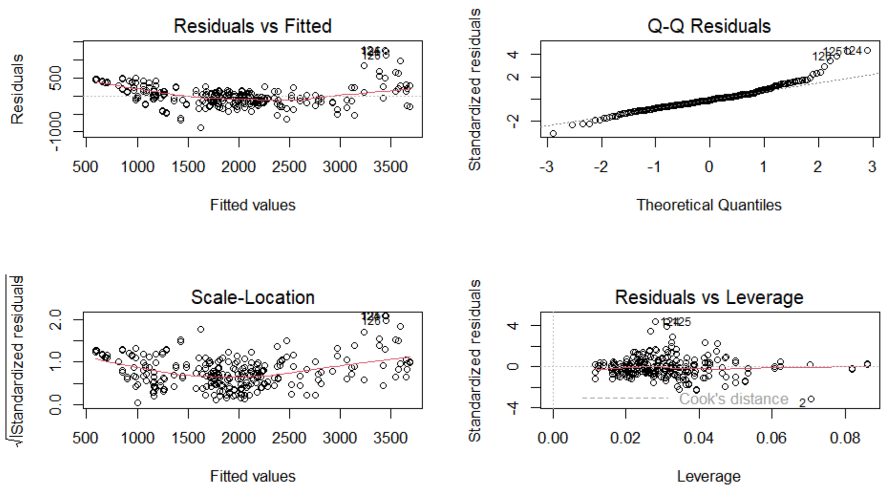

Figure 27: Result of the Model Diagnostic

b) Correlation matrix

```
# Load necessary libraries

library(reshape2)

library(ggplot2)



# Note that 'combined_data_clean' is the data frame containing the relevant
variables

# Calculate the correlation matrix using complete observations

cor_matrix <- cor(combined_data_clean[, c("IXIC.Adjusted",
"USALORSGPNOSTSAM", "UNRATE", "UMCSENT",

                                          "MVMTD027MNFRBDAL",
"CSCICP03USM665S", "CPIAUCSL", "FEDFUNDS")],

                  use = "complete.obs")



# Melt the correlation matrix for plotting

melted_cor_matrix <- melt(cor_matrix)
```

```
# Create a ggplot2 square plot of the correlation matrix with text annotations

ggplot(data = melted_cor_matrix, aes(Var1, Var2, fill = value)) +

  geom_tile(color = "white") +

  geom_text(aes(label = round(value, 2)), vjust = 1, color = "black", size = 3.5) +

  scale_fill_gradient2(low = "blue", high = "red", mid = "white",

                       midpoint = 0, limit = c(-1,1), space = "Lab",

                       name="Pearson\nCorrelation") +

  theme_minimal() +

  theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),

      axis.text.y = element_text(vjust = 1)) +

  coord_fixed()


# Save the plot to a file with square dimensions

ggsave("correlation_matrix_square_with_numbers.png", width = 8, height = 8)
```

Appendix 7: GARCH Models & Correlation Analysis: R Script (4.1.2)

a) Day-to-Day changes

aa) Script for the graph MSFT (Figure 16):

```
install.packages("rugarch")

library(rugarch)


# Fit a GARCH(1,1) model to the differenced data

spec <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)))
```

```
fit <- ugarchfit(spec, diff_data)


# Extract the conditional variances (volatility)

volatility <- sigma(fit)


# Plot the volatility

plot(volatility, type="l", main="Estimated Volatility of MSFT Daily
Changes", ylab="Volatility", xlab="Time")
```

Summary

```
summary(volatility)
```

Result:

```
    Index              volatility
 Min.   :1995-01-04   Min.   :0.0574
 1st Qu.:1997-09-29   1st Qu.:0.1550
 Median :2000-06-27   Median :0.2652
 Mean   :2000-06-30   Mean   :0.3400
 3rd Qu.:2003-04-02   3rd Qu.:0.4831
 Max.   :2005-12-30   Max.   :1.8262
```

Value of parameters

```
coef(fit)
```

Result:

```
           mu           ar1           ma1
 0.0091902819 -0.3706638616  0.3461268708
        omega        alpha1         beta1
 0.0001648583  0.0916246239  0.9073753763
```

ab) Script for the graph S&P 500 (Figure 17)

```
# Load required libraries

install.packages(c("quantmod", "rugarch"))

library(quantmod)

library(rugarch)



# Retrieve S&P 500 data from Yahoo Finance for 1995-2005

start_date <- as.Date("1995-01-01")

end_date <- as.Date("2005-12-31")

getSymbols("^GSPC", from=start_date, to=end_date)

sp500_data <- Cl(GSPC) # Extracting the closing prices



# Differentiate the S&P 500 data

diff_sp500 <- diff(sp500_data)[-1,]



# Fit a GARCH(1,1) model

spec_sp500 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder
= c(1, 1)))

fit_sp500 <- ugarchfit(spec_sp500, diff_sp500)



# Extract and print the coefficients

coefficients_sp500 <- coef(fit_sp500)

print(coefficients_sp500)
```

```
# Plot the volatility

volatility_sp500 <- sigma(fit_sp500)

plot(volatility_sp500, type="l", main="Estimated Volatility of S&P 500
Daily Changes (1995-2005)", ylab="Volatility", xlab="Time")
```

b) Daily Returns

```
# Load necessary libraries

install.packages("quantmod")

library(quantmod)



# Define the start and end dates

start_date <- as.Date("1995-01-01")

end_date <- as.Date("2005-12-31")



# Retrieve the data from Yahoo Finance

getSymbols("MSFT", from = start_date, to = end_date)

getSymbols("^GSPC", from = start_date, to = end_date) # S&P 500

getSymbols("^IXIC", from = start_date, to = end_date) # NASDAQ



# Extract adjusted prices

msft_data <- MSFT$MSFT.Adjusted

sp500_data <- GSPC$GSPC.Adjusted

nasdaq_data <- IXIC$IXIC.Adjusted



# Calculate daily returns
```

```
returns_msft <- dailyReturn(msft_data)

returns_sp500 <- dailyReturn(sp500_data)

returns_nasdaq <- dailyReturn(nasdaq_data)


# Compute GARCH(1,1) volatilities


library(rugarch)


compute_volatility <- function(data) {

  spec <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder =
c(1, 1)))

  fit <- ugarchfit(spec, data)

  return(sigma(fit))

}


vol_msft <- compute_volatility(returns_msft)

vol_sp500 <- compute_volatility(returns_sp500)

vol_nasdaq <- compute_volatility(returns_nasdaq)


# Plot the volatilities

par(mfrow=c(3,1)) # Set up the plotting area for 3 plots vertically


plot(vol_msft, type="l", col="blue", main="Microsoft Volatility")

plot(vol_sp500, type="l", col="red", main="S&P 500 Volatility")

plot(vol_nasdaq, type="l", col="green", main="NASDAQ Volatility")
```

```
# Assess correlation between the volatilities

cor(vol_msft, vol_sp500)

cor(vol_msft, vol_nasdaq)

cor(vol_sp500, vol_nasdaq)
```

Appendix 8: Correlation Matrix between Asset Classes: R Script (4.2.1)

```
# Required libraries

library(quantmod)

library(corrplot)

library(PerformanceAnalytics)

library(xts)


# Adjusted date range

start_date <- as.Date("1998-01-01")

end_date <- as.Date("2002-12-31")


# Retrieve NASDAQ Composite data

nasdaq <- getSymbols("^IXIC", from=start_date, to=end_date,
auto.assign=FALSE)

nasdaq_returns <- dailyReturn(Cl(nasdaq))


# Retrieve 10-year Treasury bond yield data

bond_yield <- getSymbols("^TNX", from=start_date, to=end_date,
auto.assign=FALSE)

bond_yield_returns <- dailyReturn(Cl(bond_yield))
```

```
# Retrieve Gold data

gold <- getSymbols("GC=F", from=start_date, to=end_date, auto.assign=FALSE)

gold_returns <- dailyReturn(Cl(gold))


# Retrieve Real Estate data using IYR as a proxy

realestate <- getSymbols("IYR", from=start_date, to=end_date,
auto.assign=FALSE)

realestate_returns <- dailyReturn(Cl(realestate))


# Retrieve International Market data using EFA as a proxy

international <- getSymbols("EFA", from=start_date, to=end_date,
auto.assign=FALSE)

international_returns <- dailyReturn(Cl(international))


# Merge all the returns data

merged_data <- merge.xts(nasdaq_returns, bond_yield_returns, gold_returns,
realestate_returns, international_returns)

colnames(merged_data) <- c("NASDAQ", "Bonds", "Gold", "Real Estate",
"International")


# Drop rows with any NA values

cleaned_data <- na.omit(merged_data)


# Compute and display the correlation matrix

correlation_matrix <- cor(cleaned_data)
```

```
corrplot(correlation_matrix, method="number", type="upper", tl.cex=0.7)
```

Appendix 9: Evolution of International Indexes and Correlation: R script (4.2.1)

a) R Script for the graph

```
# Load necessary libraries and define the date range

library(quantmod)

library(ggplot2)

start_date <- as.Date("1998-01-01")

end_date <- as.Date("2002-12-31")



# Retrieve stock data

nasdaq <- getSymbols("^IXIC", from=start_date, to=end_date,
auto.assign=FALSE)

dax <- getSymbols("^GDAXI", from=start_date, to=end_date,
auto.assign=FALSE)  # Using DAX as a proxy for Neuer Markt

kosdaq <- getSymbols("^KQ11", from=start_date, to=end_date,
auto.assign=FALSE)



# Handle missing values by linear interpolation

dax <- na.approx(dax)

kosdaq <- na.approx(kosdaq)



# Combine datasets

merged_data <- merge(Cl(nasdaq), Cl(dax), Cl(kosdaq))

colnames(merged_data) <- c("NASDAQ", "DAX", "KOSDAQ")
```

```
# Function to standardize data

standardize <- function(x) {

  (x - mean(x, na.rm = TRUE)) / sd(x, na.rm = TRUE)

}


# Standardize the closing prices

merged_data_standardized <- as.data.frame(lapply(merged_data, standardize))


# Convert to data frame for ggplot

merged_df_standardized <- data.frame(Date=index(merged_data),
merged_data_standardized)


# Plot standardized data

ggplot(merged_df_standardized, aes(x=Date)) +

  geom_line(aes(y=NASDAQ, color="NASDAQ")) +

  geom_line(aes(y=DAX, color="DAX")) +

  geom_line(aes(y=KOSDAQ, color="KOSDAQ")) +

  labs(title="Standardized Stock Indices from 1998-2002",

       y="Standardized Closing Price", x="Date",

       color="Index") +

  theme_minimal()
```

b) Correlation

```
# Load necessary libraries

library(quantmod)
```

```
# Define the date range

start_date <- as.Date("1998-01-01")

end_date <- as.Date("2002-12-31")


# Retrieve data

nasdaq <- getSymbols("^IXIC", from=start_date, to=end_date,
auto.assign=FALSE)

dax <- getSymbols("^GDAXI", from=start_date, to=end_date,
auto.assign=FALSE)  # Using DAX as a proxy for Neuer Markt

kosdaq <- getSymbols("^KQ11", from=start_date, to=end_date,
auto.assign=FALSE)


# Handle missing values by linear interpolation

dax <- na.approx(dax)

kosdaq <- na.approx(kosdaq)


# Compute daily returns

nasdaq_returns <- dailyReturn(Cl(nasdaq))

dax_returns <- dailyReturn(Cl(dax))

kosdaq_returns <- dailyReturn(Cl(kosdaq))


# Merge return datasets

returns_data <- merge(nasdaq_returns, dax_returns, kosdaq_returns)

colnames(returns_data) <- c("NASDAQ", "DAX", "KOSDAQ")
```

```
# Compute correlation matrix for the returns

correlation_matrix <- cor(returns_data, use="complete.obs")


# Print correlation matrix

print(correlation_matrix)
```

Appendix 10: Gold price vs S&P 500 Adjusted Close Price: R Script (5.1.1)

a) Graph

```
# Load Necessary Libraries

library(quantmod)

library(ggplot2)


# Retrieve S&P 500 data

start_date <- as.Date("1995-01-01")

end_date <- as.Date("2005-12-31")

getSymbols("^GSPC", src = "yahoo", from = start_date, to = end_date)

sp500_data <- data.frame(Date=index(GSPC),
GSPC.Adjusted=GSPC$GSPC.Adjusted)


# Retrieved gold data

gold_data <- data.frame(

      Date = as.Date(paste0(as.character(1995:2005),"-01-01")),

    Gold_price = c(383.79, 387.81, 331.02, 294.24, 278.98, 279.1, 271.04,
309.73, 363.38, 409.71, 444.74)

)
```

```
# Aggregate data to obtain annual closing value

sp500_annual <- aggregate(GSPC.Adjusted ~ format(Date, "%Y"), data =
sp500_data, tail, 1)

colnames(sp500_annual)[1] <- "Year"

sp500_annual$Year <- as.Date(paste0(sp500_annual$Year, "-01-01"))



# Merge data

merged_data <- merge(gold_data, sp500_annual, by.x="Date", by.y="Year",
all.x=TRUE)



# Plot data

y_range <- range(merged_data$Gold_price, merged_data$GSPC.Adjusted, na.rm =
TRUE)

plot(merged_data$Date, merged_data$Gold_price, type = "l", col = "gold",
ylim = y_range, ylab = "Price", xlab = "Date", main = "Gold vs S&P 500
(1995-2005)")

lines(merged_data$Date, merged_data$GSPC.Adjusted, col = "blue")
```

b) Correlation coefficient

```
correlation <- cor(merged_data$Gold_price, merged_data$GSPC.Adjusted, use =
"complete.obs")

print(correlation)
```

Appendix 11: Table Computation of Sharpe Ratio, Balanced Portfolio 1995-2005 (5.2.5)

Table 1: Evolution Sharpe Ratio of Balanced Portfolio, 1995-2005

Annual risk-free rate	Year	Annual Returns	Annual Standard Deviations	Sharpe Ratio per Year
0.0657	1995	0.2549245	0.07683714	2.4626697
0.0644	1996	0.18102799	0.10351461	1.12668142
0.06	1997	0.21656087	0.15318313	0.99920184
0.0526	1998	0.2424994	0.14570224	1.30333891
0.0565	1999	0.14413343	0.15735783	0.55690546
0.0603	2000	-0.09568229	0.19528783	-0.79873021
0.0502	2001	0.15833411	0.13580965	0.79621815
0.0461	2002	-0.05907146	0.15137145	-0.69479059
0.0401	2003	0.13312554	0.14606027	0.63689832
0.0427	2004	0.20621224	0.08471986	1.93003439
0.0429	2005	0.09714178	0.13780544	0.39361129