Sequential Tests for Detecting Speculative Bubbles

Niklas Michl

7406630

nmichl@smail.uni-koeln.de

Schießmauerstraße 19, 63755 Alzenau

Master's Thesis

Faculty of Econometrics and Statistics

Universität zu Köln

First Examiner: Prof. Dr. Jörg Breitung

Second Examiner: Jun. Prof. Dr. Sven Otto

Köln, 29. 04. 2025

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
AR:	Autoregressive
ACF:	Autocorrelation Function
ADF:	Augmented Dickey-Fuller Test
BIC:	Bayesian Information Criterion
CAPM:	Capital Asset Pricing Model
CDR.WA	CD Projekt SA
CGC	Canopy Growth
CUSUM:	Cumulative Sum
INTC	Intel
MSFT	Microsoft
PLUG	Plug Power
QCOM	Qualcomm
ORCL	Oracle
SADF:	Supremum Augmented Dickey-Fuller
GSADF:	Generalized Supremum Augmented Dickey-Fuller Test
VOW.DE	Volkswagen

1 Introduction

1.1 Presentation of the Topic

For as long as there have been stock markets, investors have tried to find certain trading strategies to generate high returns. Most hedge fund strategies serve as prominent examples of various investment approaches implemented to "beat the market" and generate access return. However, a bet made by one of the world's most well-known investor, Warren Buffet, clearly suggests that "beating the market" is challenging. In this bet Warren Buffet said that no one could select hedge funds which will outperform a simple investment in an index fund, namely the Vanguard S&P index fund, over a 10-year period. Ted Seides the betting counterpart invested in five hedge funds which themselves invested in over 100 hedge funds. The results have shown that none of these hedge funds outperformed the index fund and three of them delivered a return under 10% while the index fund delivered a return of 85.4% (Berkshire Hathaway Inc., 2016). Nevertheless, Warren Buffet states that there are some investors which are likely to outperform an index fund like the S&P. Even so, he admits: "In my lifetime, though, I've identified – early on – only ten or so professionals that I expected would accomplish this feat" (Berkshire Hathaway Inc., 2016, p. 24). This clearly illustrates the difficulty of outperforming the market and developing trading strategies that can lead to excess returns.

A fundamental aspect for traders is to early on identify upward and downward trends in the stock price. Many different methods, like value investing or the technical chart analysis, can be considered in order to capture these trends. Whenever a trend is identified the traders will implement a short or long position, depending on the direction of the trend, in order to gain positive returns. Dividends and short positions disregarded, its apparent that an investor can only make positive returns with long positions as long as the price of the stock is rising. In some stocks, episodes of bubbles can occur, during which the increase in price accelerates very rapidly. Therefore, it is only logical to consider exploiting these episodes to develop a trading strategy which can "beat the market". There are two main problems. The first hurdle arises from the fact that the phase of fast increasing prices is often followed by a severe downfall as seen in instances such as the Dotcom bubble around the 2000s (Phillips et al., 2011). These price patterns are depicted in figure 1.

Figure 1

NASDAQ Composite

Note. Data from Yahoo Finance (n.d.-a) and Federal Reserve Bank of St. Louis (n.d.) Retrieved March 31, 2025. See Table 2 for full details. Real prices were calculated by adjusting the NASDAQ Composite Index (^IXIC) for inflation using the Consumer Price Index (CPIAUCSL).

The second problem is the early identification of bubbles to have sufficient time to profit from price increases before the bubble bursts, driving prices back to pre-bubble levels or even lower. In the following an attempt will be made to address the aforementioned problems to some extent. Both problems will be approached by employing tests which can identify structural changes. In this context structural change represents a substantial shift in the dynamics of stock prices, for example from a random walk to an explosive process. These tests present the opportunity to invest early in bubbles and utilize the above mentioned increase in prices to gain profits. By serving as indicators, the tests will assist in determining also an appropriate selling point for the stock, although this remains particularly challenging as they were not designed for this purpose (Phillips et al., 2011). The exact techniques and the fundamental functionality of these tests will be explained later.

Despite the fact that since Phillips et al., (2011), the testing procedures have been so sophisticated that they are able to pin down the existence of a bubble early after their occurrence. The subsequent research concerned with these tests mainly focused on comparing tests to identify the best test for different scenarios and investigated them to better understand their properties (Astill et al., 2018; Homm & Breitung, 2012; Otto & Breitung, 2022; Phillips et al., 2015a). Most of these papers have the aim to only identify bubbles or try to investigate under what condition they might fail to detect them (Astill et al., 2018; Harvey et al., 2015). But none of these approaches thought about the potential to generate returns by identifying a bubble early on and investing in it. Therefore, a significant research gap presents itself, which will be addressed by analyzing various stock data to determine whether such tests can be utilized to develop advanced trading algorithms capable of generating excess returns.

The Dotcom of figure 1 is used as an example to show how structural change testing might provide excess returns. It shows the potential of realizing substantial returns if bubbles are predicted early on: "Similarly in 1998 the market value of the stocks traded in the NASDAQ increased at an annual rate of 41 percent; in the subsequent 15 months they increased at the annual rate of 101%" (Kindleberger & Aliber, 2005, p. 27). Specifically, the potential gains that may be obtained by entering the market at the start of a bubble, indicated by the structural break test, and selling the stock index before the bubble bursts, will be demonstrated.

To achieve this, the NASDAQ (^IXIC) index will be utilized, as it experienced the most severe rise in prices in all U.S. stock indices during the 2000 financial bubble (Phillips et al., 2011). The used data consists of monthly prices of the NASDAQ index. The observation period spans from the beginning of 1985 until the end of 2002. In order to transform the nominal prices into real prices the Consumer Price Index (CPIAUCSL) is used. The Supremum Augmented Dickey-Fuller test (SADF) procedure of Phillips et al. (2011) and the Generalized Supremum Augmented Dickey-Fuller test (GSADF) of Phillips et al. (2015a) are applied to determine the presence of a bubble in the data, while the Backward Supremum Augmented Dickey Fuller Test procedure of Phillips et al. (2015a) is utilized to identify the specific point in time where the bubble began to emerge.

Illustrated in table 1 are the SADF and GSADF test statistic, which clearly exceed their critical values with 6.12 for their respective test statistic at all significance level. This indicates a bubble in the NASDAQ dataset from 1985 until the end of 2001.

Table 1

Critical Values and Test Statistics

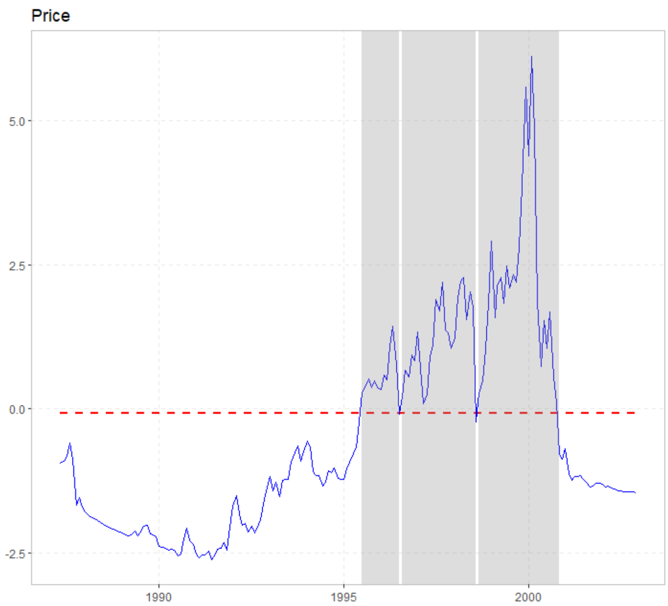
Stat	T-stat	90% CV	95% CV	99% CV
SADF	6.12	1.10	1.39	1.95
GSADF	6.12	1.86	2.11	2.58

Note. Data from Yahoo Finance (n.d.–a) and Federal Reserve Bank of St. Louis (n.d.). Retrieved March 31, 2025. See Table 1 for full details. SADF and GSADF test statistics are calculated by the author based on the resulting real price series.

The date stamping procedure in figure 2 identifies a bubble as early as July 1995 and the peak of the bubble in February 2020. Accounting the loss in purchasing power by using real prices the return would reach approximately 321.09% if an investor could identify the end of the bubble perfectly. Using the annualized geometric average return the return would be about 38.1%.

Figure 2

Date Stamping



Note. Data from Yahoo Finance (n.d.-a) and Federal Reserve Bank of St. Louis (n.d.) Retrieved March 31, 2025. See Table 2 for full details. Critical values = red dashed line. Blue line = test statistic.

This can be compared for example to the annualized geometric average return of the Russell

3000 index form January 1979 to December 2008 of 10.9% according to Christopherson et al., (2009). This illustrates the potential of conducting real-time monitoring with structural break tests. Finding an appropriate rule to invest before the bubble emerges and sell before the prices crash can lead to enormous returns exceeding by far those returns traditionally realized with passive investing. Moreover, an index is considered to be a good yardstick for the comparison of returns. It has been quite difficult for many investors to beat the returns realized through passive investing in an index (Christopherson et al., 2009). Furthermore, the risk should be moderate as bubbles can be identified very early on and so the drop height after a crash should not be that high.

1.2 Aim of the Thesis

The aim of this thesis is to develop a rule-based trading strategy that is built on existing bubble detection tests and can systematically generate excess returns. The challenge is to identify valid tests and implement their predictive power in practice for profitable trading. Given the high volatility and unpredictability of the markets, this topic is highly relevant as it could enable investors to better manage risk and gain profits from market developments like bubbles instead of falling victim to speculative excesses. This topic combines scientific methodology in the field of econometrics with practice-oriented investment applications.
The research question accordingly reads as follows: Is it possible to generate excess returns based on sequential bubble-detection tests applied to a stock containing a bubble?

1.3 Method

To solve the research question the SADF test as well as the Cumulative Sum (CUSUM) test will be considered. These tests are applied to eight stocks where bubbles have been confidently identified by detecting explosive behavior in their stock prices. Both tests operate under the same basic structure. The null hypothesis states that the whole stock price series has a unit root structure, while the alternative hypothesis assumes a shift to explosive behavior at some point. This approach is similar to many other applications of the tests, such as those in Homm and Breitung (2012) and Phillips et al. (2011). Because the possibility of a real world implementation shall be demonstrated, the chart histories are deliberately disregarded, like they were unknown. The first step is to define a training sample, which does not contain a bubble, for an initial estimation of the tests. From there on one data point will be added at a time and repeatedly the aforementioned tests will be applied to the dataset. This course of action simulates a real-time monitoring procedure. Investment will commence upon the indication of a market bubble, as identified by the specified test. Subsequent to this, the stock shall be retained until the test statistics dictate its sale by falling below the critical value again. The returns generated from these investments will subsequently be analyzed, whether they yield positive excess returns.

1.4 Structure Thesis

The first part of this thesis will summarize the literature. It starts with the definition of important key terms and lays out the theoretical framework for bubbles. Subsequently, the section on theoretical foundations explores the functionality and basic ideas behind the tests. Further the current state of research in the bubble detection field is presented. Subsequently, it is shown how the tests are implemented and adapted for the real-time monitoring framework. Moreover, how the investment performance based on the test is exactly evaluated and the data is described all in the Methodology section. In the Empirical analysis section the exact specification of the test is determined based on the training sample. Also, the results will be presented here. Next is the final conclusion with insights whether the research question could be answered. Finally, the reference list with all sources used in this thesis and the appendix are given.

2 Literature Review

This section establishes the theoretical foundation upon which this study is built. It defines key concepts, presents relevant theories, and summarizes existing research. It helps to get a sound understanding of the implemented methodologies. Specifically, it explores the SADF and CUSUM test, mainly based on their ability to detect the emergence of bubbles. Based on that the reader should get a sound understanding how these strategies lead to the returns presented in the empirical analysis section. In this way, the theoretical framework directly supports the study's goal of developing a practical investment algorithm for stock market trading through implementing these tests in a real-time monitoring scenario.

2.1 Definition of Key Terms

The first term that needs to be mentioned in this section is the standard Brownian motion, also called the "Wiener process". It is not necessary to understand what a standard Brownian motion is to comprehend the results and the approach of this thesis, nevertheless it defines the limiting distribution of both tests and therefore plays an important role in the conducted research. The standard Brownian motion is a continuous time stochastic process. Its increments can be best explained as the continuous counterpart to white noise shocks prominent in discrete-time models (Tsay, 2005). The standard Brownian motion W(r) has three key properties:

1. They have independent increments, so for any $0 \leq r_1 < r_2 < r_3$ the increments $W(r_2) - W(r_1)$ and $W(r_3) - W(r_2)$ are independent of each other.

2. For any r_1, r_2 with $0 \leq r_1 < r_2$ the increments $W(r_2) - W(r_1) \sim N(0, r_2 - r_1)$.

3. $W(r)$ has a continuous sample path.

Only the increments of a standard Brownian motion are normally distributed with zero mean and a variance of 1. Other Brownian motion may have different distributions (Campbell et al., 1997). To understand why the test statistics converge to a standard Brownian motion in their limits it is necessary to preview a part of the explanation of these tests provided further down. Both tests, the SADF and the CUSUM rely on tracking the partial sum of the residuals of the original series which are under the null hypothesis independent normally distributed. The Functional Limit Theorem or also called Donsker's Theorem states, that such a random variable (residual in this case) $u_t \sim N(0, \sigma^2)$ converges in their sequence defined by the scaled partial sum $S_t(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} u_t$ to a Brownian motion as t goes to infinity (Tsay, 2005).

The next important concept is real-time monitoring. Real-time monitoring procedures slightly differ from tests applied to fixed historical datasets. To have efficient hypothesis tests a controlled size and power is needed. The size is nothing else than the probability of a test to reject the zero hypothesis when its actually true. This is also called a type I error. The power of a test gives us the probability of correctly rejecting the zero hypothesis when the alternative hypothesis is true. Which in turn controls the type II error which is the probability that the test fails to reject the zero hypothesis and the alternative hypothesis is true (Stock & Watson, 2019). So both size and power should be at a reasonable level so the deliver reliable results. The critical value is used to control the size of the test. It does so by defining the value of the test statistic at which the hypothesis rejects the null hypothesis so the probability of the type I error is as specified.

The problem is at a real-time monitoring scenario constantly new observations becomes available. This brings new information about the time series. Therefore, the test gets applied repeatably to the time series for each new observation. Most of the time econometricians choose the type I error to be at 5% (Stock & Watson, 2019), but even with 1% in the real-time monitoring scenario the repeated testing bears a problem. Repeatedly applying the test results in an over-rejection of the null hypothesis. The problem occurring for one shot tests which are designed to be applied to a fixed sample is the so called multiple testing problem (Astill et al., 2018). To understand this problem the following example should be helpful. If a size of 1% is used and the test is applied 100 times to a time series it needs to be expected to falsely reject at least $100 * 0.01 = 1$ the null hypothesis, even it is true. To adjust for this problem the critical values/ boundaries need to change with the increasing sample size (Chu et al.,1996). How exactly this problem is solved is explained in the test implementation section.

To understand the results of this thesis it is important to understand the term Excess returns. Excess returns measure the performance of an investment but in contrast to actual returns it sets the returns of the investment in relation to the return which could be expected based on the risk and market conditions. It is a more accurate measure than actual returns to measure whether an investment strategy is outperforming the market (Chen, 2021). The formula looks like this:

$$ExcessReturn = ActualReturn - ExpectedReturn \tag{1}$$

The expected returns get calculated differently based on the approach. One of the most prominent ways to calculate the expected excess return is the Capital Asset Pricing Model (CAPM). The CAPM postulates that the return of an asset which can be expected by an investor should be equal to the market return adjusted by the sensitivity of the asset return to the market return (Christopherson et al., 2009). The formula for the expected return from equation (1), according to the CAPM, is:

$$E(R_i - R_f) = \beta_i \, E(R_m - R_f) \tag{2}$$

With R_m denoting the market portfolio return, R_i the return of the asset and R_f the risk-free rate. β_i is the factor which measures the sensitivity of the asset return to changes in the market return (systematic risk). The risk-free rate is included as it is an important alternative for investors to gain returns without risk and therefore the expected return of a risky asset should always be above the risk-free rate (Christopherson et al., 2009). The part on the right hand side of the equation accounts for the fact that every stock has a risk which is associated with the movements of the market. If the market goes down the stocks with a positive β greater than 1 are also going down but in expectation more than the average stock given by our market portfolio and vice versa. With this approach a good risk-adjusted excess return can be created the so called Jensen's alpha.

$$\alpha_i = E(R_i - R_f) - \beta_i \, E(R_m - R_f) \tag{3}$$

Which is simply the difference between the expected return postulated by the CAPM in equation (2) and the actual return. Usually, a simple index is used for the market portfolio like the S&P 500. Through the CAPM the expected return of an asset is directly related to the movements of the market during the time of real-time monitoring procedure. Therefore, it can indicate whether a trading strategy is adding value beyond what market exposure would provide. When α_i is greater than zero it implies that the return of the trading strategy was higher than that what the CAPM would have predicted based on the systematic risk given by the co-movements of the stock and the market (Christopherson et al., 2009).

Monte Carlo simulations is also a term probably not known to everyone. Explained in a very general way a Monte Carlo simulation is providing possible outcomes of a random variable by simulating its probability distribution (IBM, n.d.). It is often used to generate critical values for a test statistic, which has a complicated or unknown distribution. It is done by taking the null hypothesis given for the test and generating data under the statistical properties of this null hypothesis many times (Stock & Watson, 2019). Based on each of this dataset the test statistic is calculated and the 95% percentile is taken to get the 5% finite sample critical values. According to asymptotic theory this yields the limiting distribution of the process as the sample size tends to infinity. Of course, usually for a large enough sample size one could always use the critical values derived by the limiting distribution of the test statistic. The problem in this thesis is, also the overall sample size might be large, due to the real-time monitoring the sample size is small at the beginning. Consequently, these aforementioned critical values might not be appropriate anymore. Therefore, finite sample critical values derived from Mote Carlo simulations, which approximate the finite sample distributions help to get more accurate critical values.

Heteroskedasticity is also an important point that needs to be considered. The error term can be homoskedastic or heteroskedastic. If the error term's variance is constant for the different input values X_i, the error term is called homoskedastic. The error is heteroskedastic if his variance is changing for the different X_i (Stock & Watson, 2019). This needs to be distinguished into unconditional and conditional heteroskedasticity. Financial time series often exhibit conditional heteroskedasticity. This can be seen by the fact that days with high volatility in asset returns are most of the time followed by days with high volatility as well, so called volatility clustering. This indicates that the volatility is correlated (Campbell et al., 1997). Unconditional heteroskedasticity occurs when a time series exhibits structural breaks in the variance of the error terms which is the case if there is for example a bubble (Harvey et al., 2015). This results in a non-stable variance which can affect the limiting distributions of the test. As the application in this thesis is stock data, the variance of the error terms can sometimes be not constant and therefore unconditional heteroskedastic even without the presence of a bubble (Harvey et al., 2015). Hence, it need to be carefully considered whether this is a major problem which needs to be accounted for in the test to get reliable results in this thesis.

Moreover, autocorrelation is a term which needs to be introduced. When a times series is autocorrelated, the values today are simply correlated with it own past values (Stock & Watson, 2019). To not account for serial correlation, when it is present in the data, can lead to several problem as well. It can result in inaccurate parameter estimates and incorrect critical values which in turn lead to size distortions. How to deal with potential autocorrelation will also be further investigated below.

2.2 Theoretical Framework of Bubbles

Episodes of explosive behavior in the price of an assets or indices is observed in many instances. The reason for that can be manifold. The most dominant explanation is the existence of a rational bubble (Phillips et al., 2011). Thus, an introduction in the theory of rational bubbles and an explanation why they emerge is beneficial for the overall understanding. This theory serves as the cornerstone for both the theoretical modeling of bubbles and the empirical analysis of tests for bubble detection in many papers, such as those by Evans (1991), Homm and Breitung (2012) and Phillips et al. (2011). It provides a framework for the economic intuition behind speculative bubbles and provides a testable structure to distinguish "normal" price movements from explosive ones. Important to note is that rational bubbles are just simply a subgroup of speculative bubbles besides irrational bubbles (Kortian, 1995).

The concept of rational expectations and behavior provides a foundation to explain rational bubble behavior as deviations from the intrinsic value (kortian, 1995). A rational bubble can be explained best by the following quote: "Simply stated, a rational bubble is present whenever an asset price deviates progressively more quickly from the path dictated by its economic fundamentals. The growth of rational bubbles reflects the presence of arbitrary and self-confirming expectations about future increases in an asset's price"(Kortian, 1995, p. 7). The term "arbitrary" refers to expectations that are not based on economic fundamentals but are determined by factors such as psychological influences. The concept of "self-reinforcing expectations" refers to the phenomenon whereby the rise in the price of an asset will reinforce the belief of a further rise in the price. Thereby attracting further investment in the asset. In this context, the existence of a bubble can be considered rational.

First it is necessary to understand how prices are evolving when there is no bubble present to better separate these two price structures from one another. The standard model for stock prices, which is used in many papers, including the ones mentioned in the beginning of this subsection, is the present value model. It states that the price of an asset is determined by its expected future dividends discounted by a discount rate (Campbell et al., 1997). The next important concept here is the Efficient Markets Hypothesis. Introduced by Fama (1970), it postulates that financial markets are efficient in terms of immediately reflecting all information, which influence the intrinsic value of an asset. The assumption is that asset prices price in all available information and change only based on new information. In the present value model this information would be priced in with the expected future dividends. Persistent deviations from the intrinsic value, such as those caused by bubbles, are unsustainable and quickly corrected by rational market participants (kortian, 1995).

So arguably the Efficient Markets Hypothesis contradicts the existence of bubbles, but as aforementioned two price structures exist. On the one hand the "normal" price structures, there is no bubble present and their price patterns align with efficient markets. Blanchard and Watson (1982) show, that under rational expectations and efficient markets the fundamental price of an asset is determined by the following formula:

$$P_t = \frac{E[P_{t+1} + D_{t+1}]}{1+R} \tag{4}$$

P_t is the stock price in period t and and D_{t+1} is the dividend payed to the investor for holding the asset between t+1 and t. R is the risk-free rate which is often considered as constant in this model (Evans,1991).
Solving equation (4) through forward iteration, by inserting the formula of P_t for $P_{t+1}, P_{t+2}, ...$, yields the following solution which describes the fundamental price:

$$F_t = \sum_{j=0}^{\infty} (1+R)^{-j} E_t \; [D_{t+j}] \tag{5}$$

From the formula it is easy to see, the fundamental price is just today's value of all expected dividend payments. Imposing the transversality condition on (5), provides the unique solution $P_t = F_t$ (Homm & Breitung, 2012):

$$\lim_{k\to\infty} E_t \left[\frac{P_{t+k}}{(1+R)^k} \right] = 0 \tag{6}$$

The transversality condition in (6) ensures that the expected discounted future price of an asset approaches zero as the time horizon extends to infinity. The condition enforces that the price of an asset is tied to its fundamental value in the long run and rules out speculative bubbles. Based on the present value model it can be stated the following: "Even though the stock price P_t is not generally a martingale, it will follow a linear process with a unit root if the dividend D_t follows a linear process with a unit root" (Campbell et al., 1997, p. 257). The theory behind the fundamental price of an assets will be reflected later in the thesis in form of a random walk under the null hypothesis.

However there is no: "a priori justification" (kortian, 1995, p. 9) to impose this tansversality condition. This yields infinite solution for the price equation and therefore allows on the other hand also for price structures with a bubble (Homm & Breitung, 2012). Yielding the following solution:

$$P_t = F_t + B_t \tag{7}$$

with

$$B_t = (1+r)^{-1} E_t \, B_{t+1} \tag{8}$$

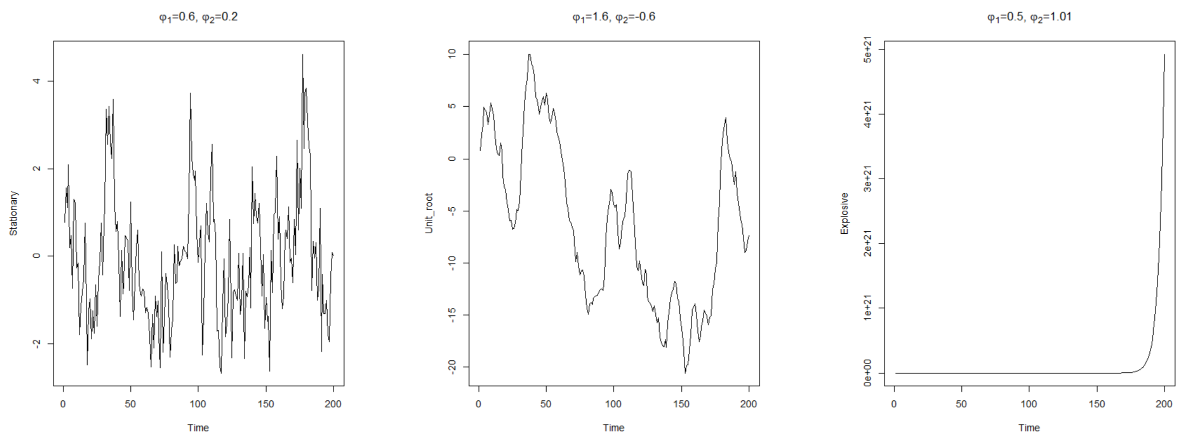
The formula (7) in combination with (8) describes the behavior of investors mentioned above. Namely that rational investors are willing to invest when such a bubble occurs, based only on the expectations that the price will go up (Homm & Breitung, 2012).

To better understand the differences in price patterns a random walk, a stationary process and an explosive process are plotted in figure 3. The random walk represents that prices evolve over time due to its shocks, which have permanent effects on the price level (Blanchard and Watson, 1982). In some cases, there may be no mean reversion, accompanied by a non-constant variance. In contrast, a stationary series fluctuates around a fixed mean and has a constant variance, as noted by Kirchgäßner et al. (2007) and shown in Figure 3(a).

Figure 3

AR(2) Processes

(a) Stationary process **(b)** Unit Root Process **(c)** Explosive Process



Note. Data is simulated by the author.

Rational bubbles can be represented by so-called mildly explosive processes, which are characterized by a root near unity that converges to unity as the sample size approaches infinity (Pavlidis et al., 2018). The last panel in Figure 3(c) illustrates such an explosive process that can be identified using structural break tests. In contrast to unit root processes, mildly explosive processes exhibit more accelerating growth patterns, with both mean and variance increasing exponentially (Kirchgäßner et al., 2007). The lack of mean reversion in explosive processes is a critical factor, which can be used to identify them. However, bubbles appear mean reverting through their burst.

2.3 Theoretical Foundation

As the equation of these tests can vary in their exact implementation, especially the CUSUM test, the first part will be introducing a basic formula of these test and implement a sound understanding why these tests are useful in the application. The tests are based on an AR(1) model like in Homm and Breitung (2012) and Phillips et al. (2011). This leads to the following hypothesis:

$$y_t = p_t y_{t-1} + \varepsilon_t \tag{9}$$

H_0: with $p = 1$ for the whole process $(t = 1, 2, ..., T)$.
H_1: with $p = 1$ for $(t = 1, 2, ..., T_e)$ and $p > 1$ for $(t = T_e + 1, ..., T)$.
With T_e denoting the emergence of a bubble and $\varepsilon_t \sim NID(0, \sigma^2)$.

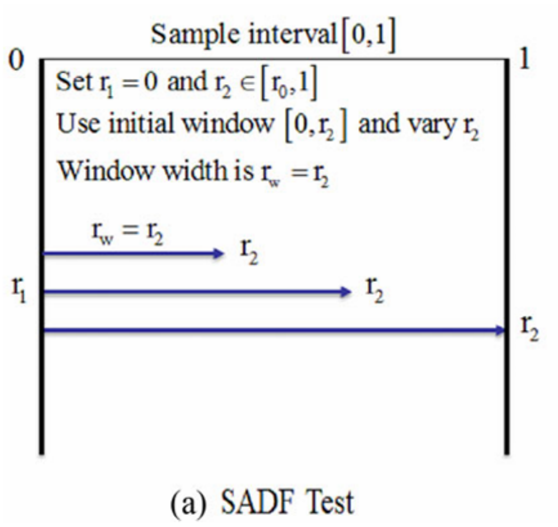
The first test to determine structural breaks in the bubble is the Supremum Augmented Dickey-Fuller test. The SADF is an extension of the Augmented Dickey Fuller test introduced by Phillips et al. (2011). It is implemented using a time series x_t to which a regular auto regressive model of the following form is applied:

$$x_t = \mu_x + \delta x_{t-1} + \sum_{j=1}^{j} \phi_j \, \Delta x_{t-j} + \varepsilon_{x,t} \tag{10}$$

The regression model is then, in contrast to the normal ADF, applied recursively. Starting from a minimum window size r_o the sample size is expanded forward by one observation for each new estimation (Phillips et al., 2011). This procedure is depicted in figure 4. As illustrated in figure 4 r_1 and r_2 are the start and the ending point of the sub period examined. The starting point r_1 is fixed to be 0 in the SADF. It provides a series of ADF tests which can be used to detect a bubble in the time series.

Figure 4

Subsamples



Source: (Phillips et al., 2015a , p.1049).

The actual test is carried out on the parameter δ. Under the null hypothesis is δ = 1 (unit root), which indicates no bubble in the time series. The alternative hypothesis would be $\delta > 1$, which indicates an explosive process in the time series and due to that a bubble could be present. The ADF test statistic is given by the following formula:

$$ADF_{r1}^{r2} = \frac{\hat{\delta}_{r_1,r_2}}{SE(\hat{\delta}_{r_1,r_2})} \tag{11}$$

To identify the most severe degree of non-stationary the supremum of the test statistics is chosen.

$$SADF(r_0) = \sup_{r_2 \in [r_0,1]} ADF_0^{r_2} \tag{12}$$

Then only this supremum test statistic is compared to the right-sided critical value and if it exceeds its critical value the test suggests a bubble in the time series (Phillips et al., 2011).

To better understand why the SADF test was developed, it is important to highlight the differences between the ADF and SADF test. As previously mentioned, the SADF test is only an extension of the ADF test. Nevertheless, the SADF test can lead to significantly better results in the detection of speculative bubbles than the classic ADF test and is one of the most applied tests to pin down explosive behavior in asset markets (Vasilopoulos et al., 2022). Usually, the ADF is used to test the hypothesis of stationary against the alternative of unit root (Stock & Watson, 2019). So when testing for bubbles the test is modified as quite the opposite it tested. The SADF compares the right-hand side critical value instead of the left-hand side critical value with the test statistic, as bubbles are characterized by a explosive time series.

Evans (1991) was one of the first to demonstrate that the ADF test has low power in detecting periodically collapsing bubbles and that these episodes of explosive behavior are often misidentified as false stationary. One intuitive explanation is that bubbles may appear mean-reverting due to their collapse (Pavlidis et al., 2018), whereas a typical explosive process that spans the entire data set does not exhibit mean-reverting behavior. However, there is also a more technical reason why the ADF test has difficulties detecting such bubbles. Evans (1991) argues that speculative bubbles are inherently non-linear, which contradicts the assumption of a linear autoregressive process implied by the ADF test. Moreover, the distribution of the test statistic resembles that of a stationary process when the probability of a bubble collapse is high. Given that bubbles cannot grow indefinitely, it is also a realistic assumption that they will eventually burst in asset markets. This illustrates the problem with the ADF test in the bubble scenario and why the design of the test needs to be adapted to successfully test for bubbles.

Therefore the design of the SADF is key to understand why it can detect bubbles better than the ADF test. The SADF identifies explosiveness in some parts of the time series through the sequence of tests. In contrast to the ADF which only examines the whole time series. To only examine explosiveness in the whole series can lead to the problem mentioned above, that the time series appears to be mean reverting. So through its design the SADF test circumvents this problem of mean reversion of an bubble as some of these ADF statistics in the sequence will exclude the collapse of the bubble.

The second famous test to detect structural breaks is the cumulative sum test also called CUSUM. The CUSUM test is accumulating deviations of an observed value from an expected value. It was developed as a refinement of the common practice to identify model variations based only on the residuals. Investigating the solely residuals is not really sensitive to changes (Brown et al., 1975). The expected value can be the mean or variance of variable of a process (Page, 1945, 1944) or deviations from a regression model (Brown et al, 1975). Since then it got further developed to be applicable also to other fields like economics, climatology and finance (Aue & Horváth, 2012). The reason for the broad usage of the CUSUM in different fields is probably it's simple construction method which can be adjusted in many ways. Nevertheless the basic framework stays the same across different implementations.

Depending on what is investigated the CUSUM formula can look quite different. A very basic formula of the CUSUM test looks like this:

$$S_t = \sum_{i=0}^{t} Y_i - \hat{Y} \tag{13}$$

$\hat{Y}$ depicts the model under the null hypothesis, which is build under the assumption of structural stability. Based on the approach this can be a linear regression model or the mean (Aue & Horváth, 2012). The cumulative sum in (13) is then evaluated at each point in time and compared with a threshold defined by a critical boundary function. If its exceeds this threshold the test suggests a deviation from the expected value or regression model defined as $\hat{Y}$, indicating a structural break in the data. The exact threshold can differ based on the context and will be further discussed in the test implementation section. The reason the CUSUM can detect structural changes such as a shift from a "normal" price behavior to a explosive one is not as obvious at first glance. When a shift occurs in the data the model under the null hypothesis is not accurately describing the underlying process. Based on the formula of the test shown above, the cumulative sum starts to deviate consistently in one direction, indicating a potential change in the model as the real values and the values from the model suggested are deviating more and more from one another. Basically, the residuals are calculated by accumulating the differences between the actual values in the time series and the expected values given by $\hat{Y}$.

The approach to calculate the CUSUM in this thesis is the same as Homm and Breitung (2012) and Breitung and Diegel (2024) use in their paper. In Homm and Breitung (2012) they use a monitoring procedure of the CUSUM approach called the CUSUM detector to sequentially detect an explosive change. They apply the following formula to calculate the Cumulative Sum.

$$S_n^t = \frac{1}{\hat{\sigma}_t} \sum_{j=n+1}^{t} (y_j - y_{j-1}) \tag{14}$$

The cumulative sum calculation starts with the first value after the training period indicated by the beginning of the sum at n+1. If sum in (14) is written out its easy to see that the formula can be simplified.

$$\sum_{j=n+1}^{t} y_j - y_{j-1} = (y_{n+1} - y_n) + (y_{n+2} - y_{n+1}) + \ldots + (y_t - y_{t-1}) \tag{15}$$

All intermediate terms in (15) cancel each other out and one ends up with the following equation:

$$S_n^t = \frac{1}{\hat{\sigma}_t}(y_t - y_n) \tag{16}$$

This approach analyzes the forecast errors directly from time series given instead of using a regression model (Homm & Breitung, 2012). Given the null hypothesis of a random walk in (9) the time series inherently lacks a clear direction as the forecast errors average out to zero. As a consequence, any deviation due to these forecast errors is purely random. The expected value for all future periods, given the value at time t, is simply y_t. This logic can simply be shown by taking the expectation to calculate what could be the one step ahead forecast. Described by the following formula:

$$E(y_{t+1}|y_t) = y_t + E(\varepsilon_{t+1}|y_t) = y_t \tag{17}$$

Under the null hypothesis ε_{t+1} is in expectation 0, so in (17) the best guess for y_{t+1} is just y_t. Therefore the formula of the CUSUM test statistic is summing all forecast errors and dividing the total by the variance of the process to evaluate a structural change. If y_t follows a random walk, the forecast errors remain small, resulting in a low value of the test statistic. If, on the other hand, it is an explosive process, y_t grows exponentially. This leads to significantly larger differences in the forecast errors and a correspondingly higher value of the test statistic, which then in turn should cross the boundary function.

2.4 Current State of Bubble Detection Research

The current bubble detection research is mainly focused on which tests are the best to identify rational bubbles in different scenarios, or in general if the test can be reliably used to identify these bubbles in the first place. Since the focus of this thesis is not on finding the overall best test for identifying bubbles, or in general explosive behavior, this section only concentrates on the research of the SADF and CUSUM test in the field of bubble detection. To assess whether both tests are suitable for the detection of bubbles one of the key points is consistency. Consistency ensures reliable results when applying the test to identify bubbles in different scenarios. As outlined in the key concepts section the rejection rate of any test, when applying it to the null hypothesis of a random walk, should be equal to the desired significance level in order to control the type I error (Stock & Watson, 2019). Phillips et al. (2015a) investigated whether the SADF has the desired significance value with asymptotic critical values for different lag length and different sample sizes. Throughout the different sample sizes and different lag lengths they identified that an overspecified lag length can lead to size distortions, but is close to the desired 5% for a small fixed lag length, which they also advocate for the empirical use of the test.

To achieve consistency also the power of the test must be accurate. Especially in this application to do not miss out on investment opportunities. However, the power of the SADF can vary a lot depending on different factors. Phillips et al. (2015a) and Homm and Breitung (2012) have shown that the power of the test increases with the number of periods examined, based on the duration and the growth rate of the bubble. Therefore, it is difficult to accurately determine whether the SADF test consistently performs well, since the aforementioned factors vary significantly in real-world applications. Yet a larger sample size can significantly increase the power of a test. The logic behind this phenomenon is the following. With higher sample size the standard errors gets better or different stated the estimates for the variables become more accurate helping also the test to become more accurate (Duflo et al., 2006).

Usually a test has reasonable power if the power is at least at 80% up to around 90% (Duflo et al., 2006). As the zero hypothesis will be that of a random walk the power is controlling in this thesis the error probability of not identifying the structural change to a bubble despite a bubble being present in the data. Since it is important to identify bubbles early on it is also reasonable to aim for a higher power around 90%. The sample size of T = 400 present as a crucial factor to reach this goal. In Homm and Breitung (2012) they report only a realized power around 10%-58%, depending also on the aforementioned factors, if the sample size is T = 100. But for a sample size T = 400 the Power level of the SADF is around 80% for most cases in their studies. Only when the bubble emerges very late in the sample or the growth rate of the bubble is very small the power is dropping below 80%. Also, Phillips et al. (2015a) report power over 80% for the SADF if the sample size is above 400 shows, that it only drops significantly below

80% if they consider very short lived bubbles. Therefore it can be concluded that the SADF test performs well in many scenarios, even if the performance is not consistently high in all cases.

Homm and Breitung (2012) and Phillips et al. (2015b) also investigated the size and power of the CUSUM detector. In Homm and Breitung (2012) the CUSUM detector has a power above 80% for a total sample size of 400 just like the SADF. If the structural break is not too late in the sample the CUSUM test achieves a power above 80% even for a smaller sample size. But for a structural change late in the sample the total sample size of 400 seems crucial for accurate power. The spike is from 0.4588% (sample size = 200) to 0.8126% (sample size = 400). Also the size seems to be accurate and close to the desired 5% (Homm & Breitung, 2012). In Phillips et al. (2015b) they also report for different bubble durations and autoregressive parameter a consistently high detection rate of the bubbles, most of the time above the desired 80%. Nevertheless, it needs to be admitted that both tests seem to have major problems when the data set is containing more than one bubble to detect each of them accurate. Efficiently only the first bubble gets detected by both tests in a multi bubble scenario (Phillips et al.,2015b). But this seems to be a minor problem in this application as the approach is to sell the stocks when there is a crash in the bubble.

It is important to note that the above mentioned power and size values are determined under the assumption of homoscedasticity as they assume normal independent errors ($\varepsilon_{x,t} \sim NID(0, \sigma^2)$) in their simulations (Homm & Breitung, 2012; Phillips et al., 2015a). As stated above conditional heteroskedasticity in the context of financial markets is a salient feature, so potential problems caused for the tests need to be considered. First of all, Phillips et al. (2011) have some empirical implementation which correctly identifies financial bubbles despite not accounting specifically for heteroskedasticity. This indicates that the test can nevertheless handle real-world volatility to some extent. Phillips et al (2015a) also examine the size of the test with conditional heteroskedasticity, which doesn't have substantial impact on the size of the test.

Nevertheless, unconditional heteroskedasticity seems to be a bigger problem. Harvey et al. (2015) report the main problem can be that the SADF is oversized which implies the test detects more often a bubble, when in fact there is no bubble present. Complementary, Astill et al. (2018) find that the size of the CUSUM test suffers when there is unconditional and conditional heteroskedasticity present in the data. In conclusion, the literature reports that only the chances of getting a signal to invest when there is actually no bubble is higher than the desired 5%. This might be a severe problem for applications in central banks where a false positive would trigger counter measurements to avoid a recession. But in this thesis an investment when no bubble has emerged appears to be a minor problem. A bigger problem would be power losses and due to that missing out on investment opportunities. Therefore, there will be no adjustment of the test used for heteroskedasticity. Another reason for that is that both tests the CUSUM and the SADF have shown being able to determine the existence of bubbles with satisfactory size and power without adjusting for heteroskedasticity.

Another point which could affect both tests is autocorrelation. Studies have shown that in financial markets especially daily return, but also weekly return data is autocorrelated to some extent (Campbell et al., 1997). However, the autocorrelation seems to be negligible according Breitung and Diegel (2024) and as Campbell et al. (1997) report that weekly autocorrelation is even smaller than daily. To neglect autocorrelation seems therefore like a feasible practice. Despite that, the autocorrelation will be additionally investigated in the implementation section by several measures to confirm the above statements. Also for the SADF Phillips et al. (2015a) advocate for the practical use a small lag length and in practice they also use a lag length of 0 to analyze the S&P 500 stock market data.

3 Methodology

3.1 Test Implementation

The implementation of the SADF test in this thesis is based on the package described in the paper Vasilopoulos et al. (2022). The regression equation is very similar to the one used in (Phillips et al., 2011) but nevertheless has some differences.

$$\Delta y_t = a_{r_1,r_2} + \gamma_{r_1,r_2} y_{t-1} + \sum_{j=1}^{k} \psi^j_{r_1,r_2} \Delta y_{t-1} + \varepsilon_t \tag{18}$$

a_{r_1,r_2} represent the intercept of the equation and is used to capture trends. This can be some kind of deterministic trend, like a linear time trend, which results in the time series to grow over time in the absence of explosiveness. Just like the formula presented in the theoretical foundation section by (10), this formula incorporates lagged differences of the time series $\sum_{j=1}^{k} \psi_{r_1,r_2}^{j} \Delta y_{t-1}$ and $\varepsilon_t \sim N(0, \sigma_{r_1,r_2}^2)$. However there are also differences which do not really change the functionality, but the hypothesis of the test. On the left side the time series is expressed in first differences with Δy_t. To derive the equation (18) based on (10) simply subtract y_{t-1} from the regression equation:

$$y_t - y_{t-1} = a_{r_1,r_2} + \delta_{r_1,r_2} y_{t-1} - y_{t-1} + \sum_{j=1}^{k} \psi_{r_1,r_2}^{j} \Delta y_{t-1} + \varepsilon_t \tag{19}$$

Simplifying (19) yields the following equation:

$$\Delta y_t = a_{r_1,r_2} + (\delta_{r_1,r_2} - 1) y_{t-1} + \sum_{j=1}^{k} \psi_{r_1,r_2}^{j} \Delta y_{t-1} + \varepsilon_t \tag{20}$$

In order to derive the formula of Vasilopoulos et al.(2022) from (20) $(\delta_{r_1,r_2} - 1)$ is set equal to γ_{r_1,r_2}. For the hypothesis test this implies that $H_0 : \gamma_{r_1,r_2} = 0$ and $H_1 : \gamma_{r_1,r_2} > 0$ is tested. This stems from the fact that the null hypothesis is that of a random walk with $\delta = 1$ and therefore $(\delta - 1) = \gamma = 0$. The effect which is used here is that by taking the first difference a process with one unit root becomes a stationary process (Hyndman & Athanasopoulos, 2021). The test statistic is now the same as in (12) with only the new coefficient γ_{r_1,r_2}:

$$ADF_{r1}^{r2} = \frac{\hat{\gamma}_{r_1,r_2}}{SE(\hat{\gamma}_{r_1,r_2})} \tag{21}$$

Which has a limit distribution given by:

$$\sup_{r_2 \in [r_0,1]} \frac{\int_0^{r_2} W \, dW}{(\int_0^{r_2} W^2)^{\frac{1}{2}}} \tag{22}$$

With W in (22) denoting a standard Brownian motion process.

However, it would be useless to apply repeatedly the SADF in such a sequential setting due to its design. At some point the test statistic would be always the same since the SADF always starts the sequence of ADF test by expanding from the minimum window size and taking the supremum. Obviously, adding in the monitoring data points after the bubble crashes would still not change the SADF statistic. The highest ADF statistic will always be the one with the window length until the bubble peak. Therefore, the test statistics sequentially used here are the

ADF test statistics. But nevertheless, the whole approach is based on the retrospective SADF, trough investigating the sequence of ADF. Due to this also the critical values of the SADF are appropriate. Additionally, the mean-reverting behavior mentioned by Pavlidis et al. (2018) can be exploited by the ADF sequence since the test statistic drops again below the critical value after the collapse. This should save the investment from plummeting in the crash phase of a bubble.

The formula for the CUSUM used in this thesis is from Breitung and Diegel (2024) which is essentially the same as (14) from Homm and Breitung (2012), but additionally divides the CUSUM by its sample size.

$$C(r) = \frac{1}{\sigma\sqrt{T}} \sum_{t=1}^{[rT]} \Delta y_t \tag{23}$$

It has a limit distribution simply given by $W(r)$, a standard Brownian motion. To additionally divide (23) through $\sqrt{T}$ is a type of normalization, which helps that the process converges to the standard Brownian motion. The estimator for the variance has the following equation.

$$\hat{\sigma}_r = \frac{1}{rT} \sum_{t=1}^{[rT]} (\Delta y_t)^2 \tag{24}$$

As noted by Breitung and Diegel (2024) this estimator in (24) can lead to size distortions when the variance is estimated for small r. Nonetheless, they report that any value of r which leads to an estimator of the variance with more than 30 observations should not exhibit this size distortions. In this application a training sample is used. The estimator therefore delivers consistent results for the monitoring period since the data of the training period is taken into account to construct the variance. As a threshold for the CUSUM test the two sided linear boundary function is used $b_\alpha(r) = \gamma_{aplha}(1+2r)$ introduced by Brown et al. (1975). γ_{alpha} is a factor which varies depending on the desired significance level. For both tests a 5% significance level is used which results in a γ_{alpha} of 0.919 for the CUSUM (Breitung & Diegel, 2024).

One aspect which seems crucial for both tests in order to get reliable results is to determine an accurate training sample size to initialize the model used. For both tests this is crucial as the law of large numbers states that under certain conditions the estimates of the parameter get more consistent for a larger sample size (Stock & Watson, 2019). Samples which are too small can generate spurious bubble signals as the coefficients of the regression are estimated based on only a few data points which can lead to finite sample biases. On the other side if the sample size is set too large the test might miss early bubbles. Additionally, Phillips et. al. (2015a) state that no bubbles is assumed to be in that training period which further emphasizes the need of a smaller training period to not include a bubble accidentally. One approach is to have around 20% of the data as training period: "[...] in practice a more realistic endpoint appears to be less than five times the size of the training sample" (Breitung & Diegel, 2024, p.28). Comparing this to the training sample size Phillips et. al. (2015a) use for the SADF and the CUSUM detector, which is the formula they also use for the SADF minimum window size $r_0 = 0.01 + \frac{1.8}{\sqrt{(T)}}$ demonstrates another approach. It varies and takes above 20% for very small samples and below 10% for very large samples. Since the sample sizes in this study vary the approach of Phillips et. al. (2015a) seems as a better fit to balance the tradeoffs explained above. Also under this rule the first estimator of the variance from the CUSUM contains already 38 observations for the smallest time series, which is more than enough to get reliable estimates for (24)

3.2 Real-Time Monitoring Framework

As mentioned above there can be an over-rejection of the null hypothesis when applying a test in a real-time monitoring scenario while the test is designed to be applied just once to the time series (Astill et al., 2018). As mentioned by Kurozumi (2020) the constant critical values can be a problem as the standard deviation may increase with an extending monitoring horizon. The CUSUM detector statistic used in this application is already designed for real-time monitoring and accounts for the aforementioned problem by implementing the time-varying boundary function $b_\alpha(r) = \gamma_{aplha}(1+2r)$.

Despite the fact that, the SADF, as constructed by Phillips et. al. (2011), is intended as a one shot test in the first place. It is also easy to implement it in a real-time monitoring framework. As explained above the SADF procedure inherently contains a sequence of test statistics that are expanded with each new observation, which is conceptually similar to the step-by-step approach used in real-time monitoring. The SADF is based on the supremum of many ADF statistics and due to this shifts its distribution under the null upward (Phillips et. al., 2011). It accounts for the multiple testing by higher critical values than the critical values of a normal

ADF test. Moreover, Kurozumi (2020) has shown that the standard deviation of the SADF is asymptotically stable for any given fixed monitoring length. In other words, the distribution of the SADF statistic converges to a limit which does not become more dispersed with the monitoring length.

This would allow to use a constant boundary derived by the limit distribution given in (22). However as mentioned above due to the small sample size in the beginning finite sample critical values will be used. They are derived by Monte Carlo simulations for the different time points in the real-time monitoring scenario using the exuber package from Vasilopoulos et al. (2022). The data under the null hypothesis, used for the Monte Carlo simulations, are generated by using numbers which are independent from one another and normally distributed with zero mean and a constant variance of 1. This mimics the behavior of the Brownian motion which then in turn defines the limiting distribution of the test statistic given above (Vasilopoulos et al., 2022).

3.3 Performance Metrics

In this evaluation any costs which normally would occur are disregard, due to transaction costs or taxes. To not take transaction costs into account is in the first place due to gradually lower costs in purchasing stocks over the last years. This is accomplished through so called neo brokers, which earn their money by getting a commission from exchange traded funds providers and trading venues. At this procedure a stock is usually acquired at a price that is slightly above its market value, while it is sold at a price that is slightly below its market value. The difference between the buying and selling price, known as the spread, is the market maker's main source of income. Part of this spread is passed on to the broker as a commission for routing the trades to them.

This procedure used here is called payment for order flow (Meyer et al., 2021). As a result, investors can trade at near zero costs with explicit cost for the neo broker Trade Republic being only 0.25% of the trade volume (Meyer et al., 2021). These costs seem rather negligible as this approach is not for daily trading and rather focuses on a longer time horizon of multiple weeks or months. Also dividends will be disregarded in this analysis as the primary focus is on returns only realized due to price movements of the stocks. Since no taxes or fees are taken into account the returns given in the result section are all gross returns.

In the section definition of key terms it was already partly mentioned that simply calculating the returns of a trading strategy and comparing it to others should not be the only measure to investigate whether it can yield excess returns. Therefore below is a more detailed analysis which measures are implemented and why they are used to investigate the performance of the trading strategy. To investigate the performance, it is important to first define how a performance is usually measured in investment strategies. An accurate definition is the following: "Performance is the return or the increase in wealth over time of an investment relative to the amount of risk the investor is taking; that is, performance measurement provides a risk-adjusted return assessment" (Christopherson et al., 2009, p.3).

Hence to investigate the performance of investment two key metrics will be needed. The first metric is the returns of the stocks, so the gain or loss calculated according to the buy and sell price of the strategies. The second component is the risk, so the volatility in the stock returns. This together accounts fo the risk-return trade-off. If one investment might lead to very high returns and the other one to very small returns there is a substantial pitfall only considering the returns and not the risk. Based on the return an investor favors the investment strategy which realized the higher return. However, it could be the case that the first investment contains stocks with higher volatility and therefore it is also far more likely to see substantial returns than in the second investment. But the higher volatility in these stocks can lead then likewise to higher negative returns.

In general, the rule of thumb is, the more risk is associated with an investment the higher should be its return in the end. This ultimately, is the reason why returns do not necessarily compensate for the elevated volatility. Stressing the necessity for a risk-adjusted evaluation to really investigate the desirability of the return compared to other investments (Christopherson et al., 2009).

With regards to these considerations the aforementioned CAPM and Sharpe ratio are implemented as primary measures beside the returns. The CAPM can yield a sound measure for the risk-adjusted returns based on the market environment. Another measure implemented and compared to the returns is the buy-and-hold strategy where it is assumed that the investor simply holds the stock for the whole monitoring period. There are two reasons to also take such a non risk-adjusted measure into account. First the performance can be compared to the experiment Warren Buffet made, mentioned in the introduction. Second it can be investigated whether the tests are able to capitalize on their strength to filter out these time points when the stocks make huge returns while avoiding getting caught in the bubble's collapse. Also, in general it is a popular approach to compare the returns to indices like the S&P 500 or the NASDAQ not least because of investing passive in these indices can lead to a good performance (Christopherson et al., 2009).

The Sharpe ratio is also a risk-adjusted measure but compared to the Jensen's alpha and the CAPM it is based on the total risk of the asset and not only the systematic risk (Christopherson et al., 2009).

$$Sharpe\ Ratio = \frac{R_i - R_f}{\sigma_i} \tag{25}$$

It compares the excess return $R_i - R_f$ to the volatility of the asset σ_i. The higher the Sharpe ratio the more efficient was the return based on the risk taken with an investment.

3.4 Data Description

In table 2 are datasets displayed. This is done to clarify which dataset underlies each figure or metric and to avoid overloading the thesis with citations. Each of these datasets is mentioned in table 2 with an acronym which will be used in tables or the text to refer to the metrics developed by the corresponding dataset.

Table 2

Data Sources for Datasets

Acronym	Source	Retrieval Date	URL
CPIAUCSL	Federal Reserve Bank of St. Louis (n.d.)	March 31, 2025	https://fred.stlouisfed.org/series/CPIAUCSL
^IXIC	Yahoo Finance (n.d.-a)	March 31, 2025	https://finance.yahoo.com/quote/^IXIC/history/
^GSPC	Yahoo Finance (n.d.-b)	March 31, 2025	https://finance.yahoo.com/quote/^GSPC/history/
^TNX	Yahoo Finance (n.d.-c)	March 31, 2025	https://finance.yahoo.com/quote/^TNX/history/
MSFT	Yahoo Finance (n.d.-d)	March 31, 2025	https://finance.yahoo.com/quote/MSFT/history/
INTC	Yahoo Finance (n.d.-e)	March 31, 2025	https://finance.yahoo.com/quote/INTC/history/
QCOM	Yahoo Finance (n.d.-f)	March 31, 2025	https://finance.yahoo.com/quote/QCOM/history/
ORCL	Yahoo Finance (n.d.-g)	March 31, 2025	https://finance.yahoo.com/quote/ORCL/history/
CDR.WA	Yahoo Finance (n.d.-h)	March 31, 2025	https://finance.yahoo.com/quote/CDR.WA/history/
PLUG	Yahoo Finance (n.d.-i)	March 31, 2025	https://finance.yahoo.com/quote/PLUG/history/
VOW.DE	Yahoo Finance (n.d.-j)	March 31, 2025	https://finance.yahoo.com/quote/VOW.DE/history/
CGC	Yahoo Finance (n.d.-k)	March 31, 2025	https://finance.yahoo.com/quote/CGC/history/

Note. Source citations correspond to Reference List entries (n.d.-a through n.d.-k and Federal Reserve Bank of St. Louis, n.d.).

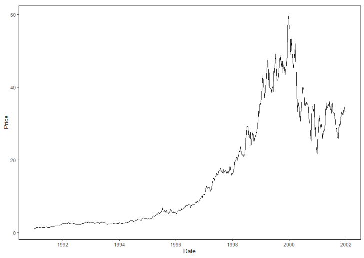
The data used in the empirical analysis is compromised of historical weekly closing price data of eight companies, the S&P 500 and the 10-Year Treasury bond yield of the United States of America . The Weekly price data of the S&P 500 (^GSPC), from 1990-01-01 until 2023-01-01, is used as the benchmark index in the CAPM model and the 10-Year Treasury bond yield

(^TNX) for the same time span as the proxy for the risk-free rate. The data is retrieved trough an application programming interface provided by the R package yahoofinancer. The companies investigated in this study are CD Projekt SA, Microsoft Corporation, Intel Corporation, Volkswagen AG, Qualcomm Incorporated, Oracle Corporation, Plug Power Incorporated and Canopy Growth Corporation depicted in figure 5.

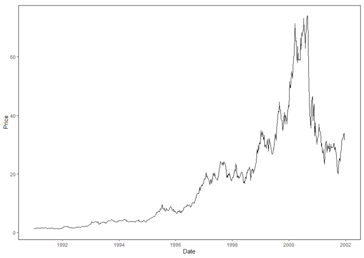
Figure 5

Stock Prices

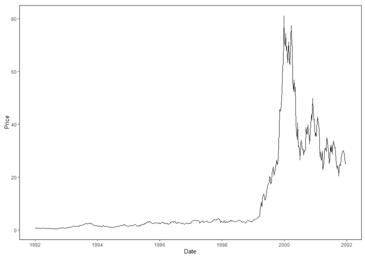
(a) MSFT



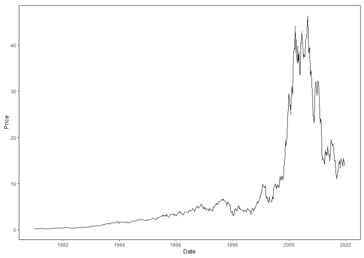
(b) INTC



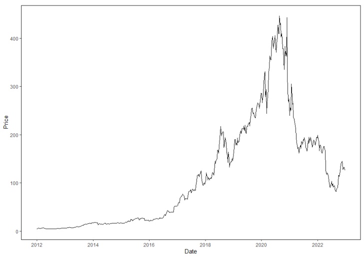
(c) QCOM



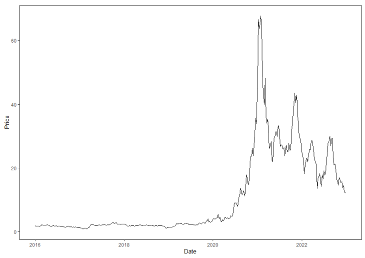
(d) ORCL



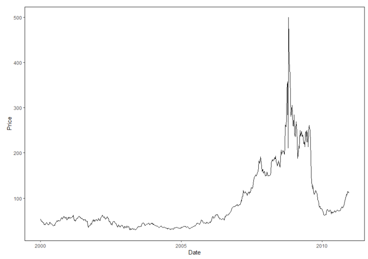
(e) CDR.WA



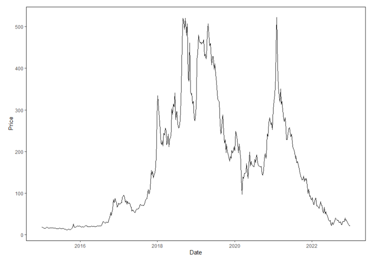
(f) PLUG



(g) VOW.DE



(h) CGC



Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details.

All these companies experienced at some point in their investigated time span an explosive behavior in their prices. Some of them are investigated in times of famous bubbles in the overall market just like Microsoft (MSFT) , Intel (INTC) and Oracle (ORCL) in the aforementioned 2000s Dotcom bubble. The observation period is ranging for these stocks from 1991-01-01 until 2002-01-01 with 547 weekly observations. Another stock which was influenced from the Dotcom bubble is QUALCOMM (QCOM = 522 observations) investigated in the time span from 1992-01-01 until 2002-01-01. Other stocks investigated in this thesis experienced very specific reasons for fast accelerating prices just like CD Projekt S.A. (CDR.WA = 547 observations from 2012-01-01, 2022-12-28), which became Polands most valuable company during the COVID-19 pandemic (Hellrung, 2020). This price pattern occurred probably due to CD Projekt S.A. being a video game company and therefore not being negatively influenced by ramifications like lockdown or supply chains distributions in contrast to many other companies during the COVID-19 pandemic.

Around 2020 Plug Power (PLUG = 366 observations from 2016-01-01 until 2023-01-01) also had fast accelerating stock prices as the pledges of green hydrogen in the upcoming years fueled the expectation of future profits in this industry (Breitung & Diegel, 2024). A company which experienced fast accelerating prices as well, due to expectations of substantial future returns in their sector is Canopy Growth Corporation (CGC = 418 observations from 2015-01-01 until 2023-01-01). These expectations have been justified by investors in prospect to a legalization of cannabis in Canada around 2018 (MacDonald, 2024). At last the stock prices of Volkswagen (VOW.DE = 547 observations from 2000-01-01 until 2010-12-24) are investigated. More specifically the stock prices of stocks with voting rights, which experienced a very sharp rise during the attempted takeover from Porsche AG in 2008 (Grah, 2019) are used. The SADF and the CUSUM tests are then applied to the original price series. The tests can and are also done by taking the logarithm of the price series by other researchers like Homm and Breitung (2012). However as Phillips et al. (2011) report that the results for the SADF are similar when using raw level price data or logarithmic price data, the tests will be made with raw level price data. All data visualizations and statistical analyses were performed in R (version 4.3.0). The used packages are ggplot2 for plotting, exuber for the SADF test and for data manipulation dplyr, lubridate and PerformanceAnalytics.

All aforementioned information considered all of these stocks experienced sharp rises in their prices due to events that aren't necessarily tied to changes in their fundamental value. Rather the increases are fueled only by expectations that the price will go up for the aforementioned reasons. However in studies like Phillip et al. (2011) they also investigate the development of the dividends over time to confirm the theoretical framework of a rational bubble mentioned above. This extra step will not be made in this thesis. For the success of the investment it is rather irrelevant why a explosive process occurs since the test is only based on the patterns of stock prices. The non-necessity of the underlying theoretical framework has also been mentioned in the context of house and commodity prices, which aren't compatible with dividend based frameworks: "Nevertheless, detecting a change from I(1) to explosive clearly points to excessive speculation" (Homm & Breitung, 2012, p. 225). Another point is that there are many other definitions of a bubble like the one of Kindleberger and Aliber(2005) describing a bubble far more simple as just a phase where prices increase very fast followed by a collapse of the bubble. Therefore it is hard to pin down an exact definition and access whether these price patters can now be defined as bubbles. To solve this question the research seems to be too extensive for a topic which doesn't help to answer the research question and would distract from the main examination of this thesis.

4 Empirical Analysis

4.1 Training Sample

As mentioned by Phillips et al. (2015a) it is assumed that no bubble occurs in the training period. Also Astill et al. (2021) emphasized that the CUSUM procedure could be harmed by a bubble in the training period. Therefore the first step is to assess whether the training periods used have indeed no bubble included. The indication of a bubble will be done by the retrospective SADF and GSADF tests. Also employing the GSADF to determine whether there is a bubble in the training period has a straightforward reason. The GSADF is not fixing the starting point. This generates more subsamples than the SADF and therefore the GSADF has power gains especially in smaller samples compared to the SADF (Phillips et al., 2015a). Since it is important to rule out the existence of a bubble, the application of both tests helps to reliably identify bubbles in the training sample. In table 3 it is evident that the SADF and GSADF test statistics are smaller than every critical value even the 90% critical value.

Table 3

SADF and GSADF Test Statistics with Critical Values

Stock	Statistic	t-stat	90% CV	95% CV	99% CV
MSFT	SADF	–1.17	0.850	1.18	1.89
	GSADF	–0.134	1.53	1.89	2.61
INTC	SADF	–1.59	0.850	1.18	1.86
	GSADF	–0.599	1.53	1.89	2.61
QCOM	SADF	–1.27	0.834	1.16	1.86
	GSADF	0.391	1.51	1.85	2.56
ORCL	SADF	0.104	0.850	1.18	1.86
	GSADF	0.820	1.53	1.89	2.61
CDR.WA	SADF	–0.690	0.850	1.18	1.86
	GSADF	0.717	1.53	1.89	2.64
PLUG	SADF	–1.06	0.826	1.16	1.81
	GSADF	–0.0483	1.44	1.82	2.56
VOW.DE	SADF	–0.499	0.850	1.18	1.89
	GSADF	0.395	1.53	1.89	2.61
CGC	SADF	–0.699	0.832	1.17	1.88
	GSADF	0.714	1.48	1.83	2.61

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details. SADF and GSADF test statistics are calculated by the author based on weekly raw price data.

For non of the training samples the test statistic is even close to the 90% critical value, which is indicating that the training samples do not show any signs of explosiveness, not even a weak level. Therefore it can be concluded that there is no statistically significant bubble at the loosest conventional significance level of 10% .

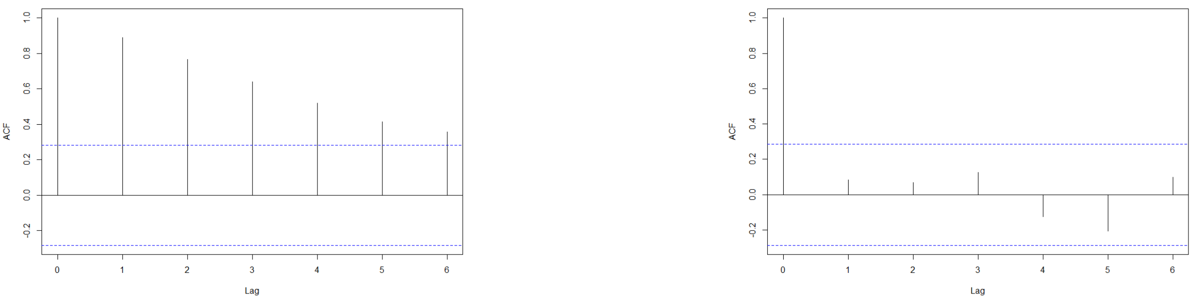
Due to the nature of this application only the training period is available to the researcher at the start of the monitoring. Therefore based on the training sample the lag length, incorporated in the test procedures, is determined. To reliably investigate whether the models presented above don't need to account for autocorrelation several steps are necessary. First to get a broad picture about the autocorrelation in the price data an autocorrelation function (ACF) plot is examined. In figure 6 two ACF plots of the CD Projekt S.A. training sample are depicted. However the plots of the other stocks looking pretty much alike and are in the appendix section figure 9.

It appears, investigating figure 6(a), the original prices exhibit substantial autocorrelation.

Figure 6

ACF Plots CD Projekt S.A.

(a) ACF of original Series **(b)** ACF of difference Series



Note. Data from Yahoo Finance (n.d.–h). Retrieved March 31, 2025. See Table 2 for full details. The ACF plots are calculated by the author based on weekly price data.

This is also confirmed by the Lung-Box test in table 4 in the original data column. The Lung-Box test is done with 10 lags as advocated by Hyndman and Athanasopoulos (2021) for non seasonal data. Every p-value of the original data series is lower than 1%, indicating very significant autocorrelation in the original price data series. But to get a better understanding which lag length would be accurate for the tests the Akaike's Information Criterion or Bayesian Information Criterion (BIC) are well established metrics (Stock & Watson, 2019). Phillips et al. (2015a) show that the lag order selection of the BIC is accurate to determine the lag length for the SADF. Therefore for every training period of the stock prices an AR model is estimated which minimizes the BIC with a upper bound of lags at 6. The BIC balances the trade off between too few lags leading to missing out on potentially valuable information and too many lags which can result in overfitting. This balancing is done by an increase for every lag in the second part of the equation $(p+1)\frac{ln(T)}{T}$ and usually a decrease in the SSE = sum of squared residuals by the estimation of a better fitting autoregressive model (Stock & Watson, 2019).

$$BIC(p) = ln\left(\frac{SSR(p)}{T}\right) + (p+1)\frac{ln(T)}{T} \tag{26}$$

As illustrated in table 4 all the stock price time series indicate a lag length of one being optimal to minimize the Bayesian Information Criterion. Therefore an AR(1) model would be accurate for the price data.

Table 4

Ljung–Box Test and Lag Selection

	P-Value		BIC	
	Original	**Difference**	**Original**	**Difference**
MSFT	$p < .001$	0.6172	1	1
INTC	$p < .001$	0.8689	1	0
QCOM	$p < .001$	0.9842	1	0
ORCL	$p < .001$	0.8411	1	0
CDR.WA	$p < .001$	0.6175	1	0
PLUG	$p < .001$	0.8396	1	0
VOW.DE	$p < .001$	0.9693	1	0
CGC	$p < .001$	0.5009	1	0

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details. Ljung-Box test statistics and Bayesian Information Criterion (BIC) values are calculated by the author based on weekly price data.

However if taking a closer look how the SADF test in (18) and the CUSUM in (23) are implemented it can be seen that both use differencing of the raw price data in their test. Differencing is often used to remove trends. differencing is making the pure random walked, which is assumed under the null, stationary and yield a white noise series (Hyndman & Athanasopoulos, 2021). Of course in practice a pure random walk is not always given but it could nevertheless help to eliminate the autocorrelation.

Trough differencing the slowly decaying bars in the ACF plot in figure (6a), which indicate a trend, are vanishing in figure (6b). Important to note is that the lag at zero is always correlated at a magnitude of 1 as it is simply the correlation of the value with itself. However also in the other ACF plots the first differencing completely removes the persistent autocorrelation presented in the plots. To reliably review whether the autocorrelation is captured already by the model design the Lung-Box test and the BIC of the differenced series is depicted in table 4 as well. The Lung-Box test for the differenced series confirms that no autocorrelation is left in the training sample with very high p-values exceeding all by far the p-value of 10%. Additionally, the BIC suggest for the differenced series a lag length of zero for most stock price data.

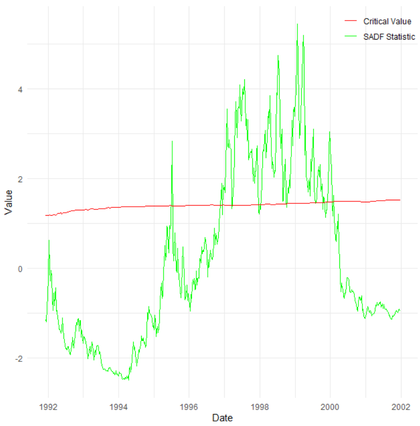
Nevertheless it needs to be mentioned that at the differenced series there is one particular outlier, MSFT. The BIC is suggesting for MSFT a lag length of 1 in the differenced series. Despite differencing wiped out the unit root, the price series of MSFT still exhibits enough first-order autocorrelation to justify an AR(1). This is however not a sign the differencing failed, but rather that some residuals have still serial dependence like for example trough volatility clustering. Nonetheless, lag order selection is in general used to determine a model which captures persistent patterns that are also robust in future periods. Conversely overfitting could lead to a greater bias in the forecast (Stock & Watson, 2019). Since 87,5% of the assets advocate a lag length of zero it is reasonable to use also a lag length of zero for the implementation of both tests. It seems to be a good fitting model to describe the usual patterns of these weekly stocks under the null while it avoids fitting to idiosyncratic noise, providing a better foundation for generalizable modeling. Therefore the autocorrelation is already captured by the designs of the SADF and CUSUM test used in this thesis and both tests can be executed with zero lag length.

After ensuring all parameters of the tests are correctly implemented it is time to analyze the results the SADF and CUSUM are providing for the different stock price series. To not overload the reader with all the results at once, as it may become very cumbersome with the results of eight stocks with two tests, the results of the SADF will be presented first and afterwards the results of the CUSUM. As a recap, the specifications used to implement the SADF are the following. The SADF equation of (18) will be applied with zero lag length, with the minimum window size set according to the rule $r_0 = 0.01 + \frac{1.8}{\sqrt{(T)}}$ and the test statistic is compared to the 95% finite critical values derived through a Monte Carlo simulation with 2000 repetitions. Figure 7 gives a first impression at which point the SADF test is indicating explosive price patterns and how long they last by depicting critical values against the SADF test statistics for four stocks. The remaining plots are illustrated in the appendix figure 10. In the plots the training period is of course left out since these periods will not be included in the evaluation of the trading rule. Investigating figure 7 it is obvious that at first glance in all four stocks bubbles are detected. Nonetheless another interesting pattern can be seen. The test statistic surpassing the critical value, indicating episodes of explosive price patterns, often drops below that critical value again in a very short period.

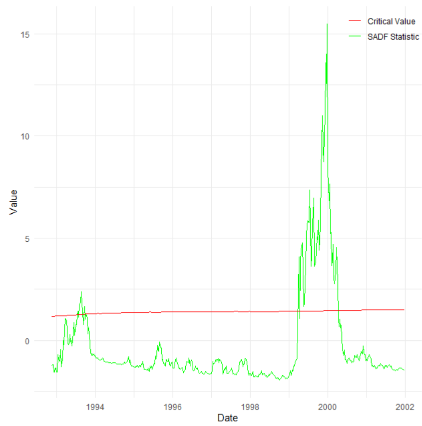
Figure 7

Critical Values vs. Test Statistic (SADF)

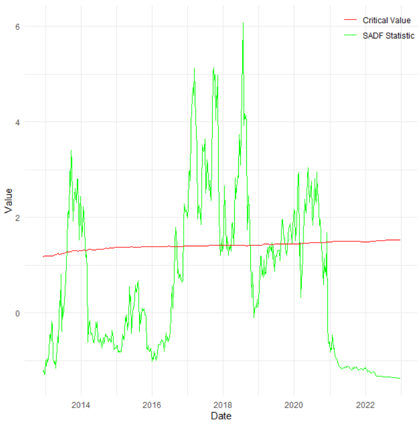
(a) MSFT



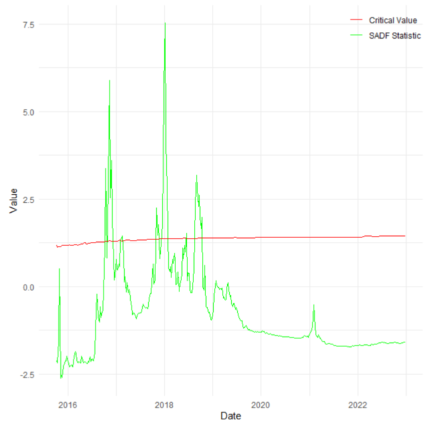
(b) QCOM



(c) CDR.WA



(d) CGC



Note. Data from Yahoo Finance (n.d.–d; n.d.–f; n.d.–h; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details. SADF test statistics is calculated directly from weekly raw price data. Depicted are here the 95% finite sample critical values calculated trough Monte Carlo simulations. Both calculated by the author.

One example for this the MSFT plot in figure (7a) where the test statistic is surpassing the critical value on 1996-11-26 and just two weeks later in 1996-12-10 it is again below that critical value for one week to again indicating accelerating price patterns in the subsequent week. Especially to have only one week where the test is falling below the critical value while subsequent and previous periods are above the critical value can be seen for nearly every stock.

It may be puzzling that the test is indicating for some stocks explosive behavior way before the time point a pure eyeball inspection of the stock price data in figure 7 would suggest. For instance comparing the stock price pattern of CDR.WA in figure(5e) with the plot in figure (7c) makes the following evident. It is reasonable to argue that the stock price seems to rise very sharply around 2016 and before that it assembles more a random walk rather than an explosive process.
Nevertheless the test statistic is detecting explosive price patterns already in 2014. However this is not indicating the test is failing, the problem is the graphs in figure 5 are in absolute terms. Due to the scale in absolute terms the price changes in the beginning look far smaller than the changes in the prices later on straightforward since the whole scale is stretched out to fit higher values later in the sample. Therefore the graphs are reflecting more accurate the price changes in the end of the sample than in the beginning. Which in turn explains why the tests detect explosive episodes while the a pure eyeball inspection fails to do that.

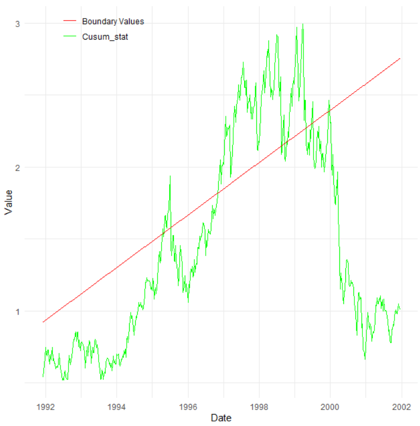
To avoid that logarithmic price series could be used as they compress the scale and make changes in the stock price more comparable since the scale is not in absolute terms anymore. It might arise the question why figure 5 is not done in logarithmic terms then. To demonstrate more clearly the connections between the events and the probably explosive price processes the original price series seems just more accessible.

Next the CUSUM test is implemented on the eight stocks. Based on the investigation above it is also reasonable to use here a lag length of zero and the critical values are based on the boundary $b_\alpha(r) = \gamma_{aplha}(1+2r)$ with the $\gamma_{5\%} = 0.919$. In figure 8 are the same four stocks depicted with the critical values against their CUSUM test statistic and the rest of the plots are in the appendix in figure 11. It can be seen that in all stocks bubbles are detected. However in figure 8 (d) and in figure 8 (b) the CUSUM detects explosiveness in only a few periods. It is directly obvious that both tests will yield very different results for some of the stocks.

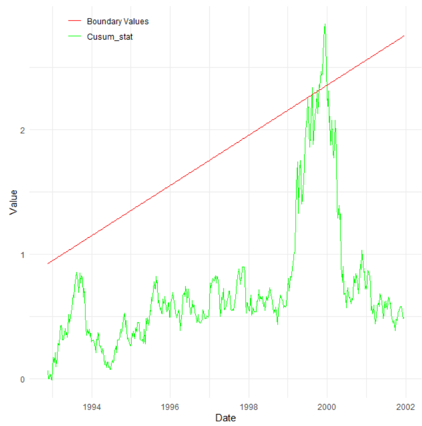
Figure 8

Critical Values vs. Test Statistic (CUSUM)

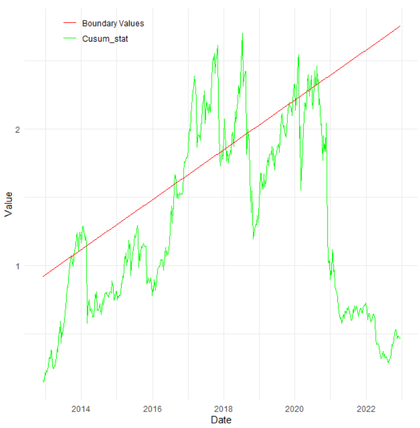
(a) MSFT



(b) QCOM



(c) CDR.WA



(d) CGC

Note. Data from Yahoo Finance (n.d.–d; n.d.–f; n.d.–h; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details. CUSUM test statistics were calculated by the author directly from weekly raw price data and compared against the derived critical boundary function with the corresponding 95% critical value for γ_{aplha}.

The test statistic patterns look quite similar with peaks and valleys occurring around the same time point with for example QCOM having the highest test statistic both at the end of 2019 in figure 7 (b) and 8(b). While the critical boundary function of the CUSUM and the finite sample critical values of the SADF have quite different patterns

4.2 Results

The trading rule is completely determined by the test. The stock is bought when the corresponding test statistic is above the critical value and sold as soon as it drops below the critical value. Therefore it is implemented in the code that the stock is bought directly at the price when the corresponding test statistic surpasses the critical value. Then the stock is hold until the test statistic is falling again below the critical value. The stock is sold at the price the first time the test statistic is below the critical value, which should indicate the transition of the stock from explosive price patterns back to a random walk again. However due to the sometimes excessive variations in the test statistics described above there are many buy and selling periods over the monitoring period.

The first metric assesses whether the trading strategy can outperform the expected excess return under the CAPM. Usually the CAPM is applied to portfolios, which diversified the idiosyncratic (firm-specific) risk of the stocks away. Consequently, the CAPM assumes the only risk of the investment is related to movements in the market which can't be diversified (Christopherson et al., 2009). This influences the estimates for the CAPM coefficient displayed in table 5.

Table 5

CAPM Regression Results via Jensen's Alpha (SADF Test)

Stock	Jensen's Alpha	p-value	Beta	p-value	R^2
MSFT	0.0019	0.1188	0.4050	< .01	0.0901
INTC	0.0004	0.6714	0.3029	< .01	0.0712
QCOM	0.0030	0.0769	0.3113	< .01	0.0339
ORCL	0.0011	0.3843	0.2967	< .01	0.0483
CDR.WA	0.0023	0.1117	0.0292	0.6433	0.0004
PLUG	0.0042	0.1256	0.0351	0.7381	0.0003
VOW.DE	0.0003	0.8050	–0.2065	< .01	0.0374
CGC	0.0006	0.7842	0.0728	0.4113	0.0018

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k; n.d.–b; n.d.–c). Retrieved March 31, 2025. See Table 2 for full details. Estimates and R^2 rounded to four decimal places; p-values below 0.01 reported as < .01.

Throughout all estimates the beta is very low ranging from 0.4050 for MSFT to -0.2065 for VOW.DE. However not all of these beta estimates are significant. CDR.WA, Plug and CGC have way to high p-values to analyze the beta. The remaining betas are statistically significant with a p-value smaller 1%. The small positive betas indicate that in most of the cases the investment strategy is moving with the market. But these co-movements with the market are much smaller than the movements in the market. Also the R^2 is very low for all investment with the highest being 0.0901. Some of the variance in the returns can typically be explained by other factors like small stocks and stocks with a high book to market ratio having higher expected returns (Fama & French, 2004). But in general it is common for stocks that most of their risk is idiosyncratic so a fairly low R^2 is considered normal (Bodie et al., 2005). However as they are across the board very low and sometimes extremely low like for CDR.WA or PLUG with 0.0004 and 0.0003, indicates that the risk of the investment strategy is not related to market movements or only to a very small extent. The Jensen's alphas are all positive but all statistically insignificant with a p-value greater than 5%. Therefore it can be concluded that the investment strategy based on the SADF could not yield returns above what could be expected given their correlation with the market.

The same can be said for the CAPM regressions using the CUSUM test depicted in table 6.

Table 6

CAPM Regression Results via Jensen's Alpha (CUSUM Test)

Stock	**Jensen's Alpha**	**p-value**	**Beta**	**p-value**	R^2
MSFT	0.0040	< .01	0.2851	< .01	0.0608
INTC	0.0025	< .01	0.1428	< .01	0.0304
QCOM	0.0029	0.0125	0.0598	0.2545	0.0027
ORCL	0.0018	0.0079	0.0883	< .01	0.0148
CDR.WA	0.0050	< .01	0.0304	0.5546	0.0007
PLUG	0.0005	0.3329	0.0080	0.6818	0.0005
VOW.DE	0.0001	0.3656	–0.0222	< .01	0.0879
CGC	0.0004	0.3502	0.0168	0.3412	0.0024

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k; n.d.–b; n.d.–c). Retrieved March 31, 2025. See Table 2 for full details. Weekly price data is transformed into return series before conducting the CAPM regressions on the whole monitoring period by the author. Estimates and R^2 rounded to four decimal places; p-values below .01 reported as < .01.

The beta values are even lower with again the investment in the MSFT stock having the highest beta with 0.2851 and the lowest beta at the VOW.DE investment with -0.0222. Also the R^2

are fairly low again. However one important difference is noticeable in table 6. There are now five Jensen's alphas which are statistically different from zero at the 5% level. Tough the beta value of the investments in CDR.WA and QCOM have very high p-values, indicating these investments are not significantly correlated to the movement in the markets. Due to this the alpha should be interpreted with caution since it absorbs this bias of beta (Jensen et al., 1972). Nevertheless this still leaves three investments with a significant Jensen alpha. The investment in MSFT yields yearly 20.8% excess return over what the CAPM would predict, INTC = 13% and ORCL = 9.36%. Therefore only 3 out of 16 investment could provide statistically significant abnormal returns under the CAPM. However there are a few shortcomings with the CAPM when applied to this investment strategy. It is important to keep in mind that these are not the CAPM of the stock rather the CAPM of the investments based on the trading rule in these stocks. Therefore sections are included in the CAPM analysis where the strategy is out of the market. In these sections the strategy yield only zero excess returns. This is the case as investors can always invest their money in bonds returning the risk-free rate yielding zero excess return in equation (1).

$$\hat{\beta} = \frac{\sum_{t=1}^{T}(R_{m,t} - \bar{R}_{m,t})(R_{i,t} - \bar{R}_{i,t})}{\sum_{t=1}^{T}(R_{m,t} - \bar{R}_{m,t})^2} \tag{27}$$

Investigating the OLS formula above makes the regression results more understandable. Through adding many zeros in (27) significantly lowers the beta. The economic reasons for the low beta is the risk-free rate has no correlation with the market and the mathematical reason is the nominator of (27) becomes zero.

Likewise, R^2 will drop very low. The R^2 provides a key figure to assess how well the variation in the asset return is explained by the variation of the market return. It helps to assess how well the regression line fits the data. However there is no market risk involved when staying out of the market. Due to that the alpha automatically inflates artificially and distort the results (Christopherson et al., 2009). Therefore also the three positive abnormal returns in the CUSUM investment approach may be too high estimates.

Given the logic of the CAPM, only considering market risk, seems to be almost the exact opposite train of thought to the investment strategy used in this thesis. Due to that there are only two conclusions that can be made here. First the variation in the strategy returns is only very low related to the return movements of the market. Second a measure like the Sharpe ratio which is a combination of the market risk and the idiosyncratic risk should nevertheless lead to reliable risk-adjusted estimates.

The subsequent metrics are compared to different investment strategies epitomized by hedge funds but also to indices. In the introduction it has been shown that it is quite challenging for hedge funds to beat a simple index. But also other strategies seem to have a hard time beating the indices. In only 23.51% of the cases actively managed mutual funds were able to outperform S&P 500 over five years (Bodie et al., 2005). The Sharpe ratios for the different investments based on the SADF and CUSUM are given in table 7.

Table 7

Sharpe Ratios for the Investment based on SADF and CUSUM

	SADF	**CUSUM**
	Sharpe Ratio	**Sharpe Ratio**
MSFT	0.5934	1.2462
INTC	0.2385	1.0747
QCOM	0.6519	0.8516
ORCL	0.3581	0.8845
CDR.WA	0.5145	1.3568
PLUG	0.6233	0.3994
VOW.DE	0.1038	0.3153
CGC	0.1220	0.3719

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k; n.d.–c). Retrieved March 31, 2025. See Table 2 for full details. Weekly price data is transformed into return series before calculating the weekly Sharpe ratios by the author. Sharpe ratios are annualized according to ($Sharp_{weekly} * \sqrt{52}$). Values rounded to four decimal places.

Christopherson et al. (2009) report by simply holding the S&P 500 between September 1987 to December 2006 the Sharpe ratio was at 0.59 and nearly the same for the Russell 3000 with 0.6. Comparing the Sharpe ratios in table 7 with a passive investment in these indices is therefore the first approach to assess whether the risk taken is justified by the return. For the investment approach using the SADF test only two stocks yield a Sharpe ratio slightly above the indices, QCOM with 0.6519 and PLUG with 0.6233. Also MSFT appears to perform at least similarly

well with a Sharpe ratio of 0.5934. All others investments yield a Sharpe ratio far below the Sharpe ratio of the indices. Investigating the Sharpe ratios of the investment using the CUSUM approach in table 7 provide better results. Nearly all Sharpe ratios are higher than their counterparts, using the SADF, with three stocks exhibiting a Sharpe ratio greater than one, namely MSFT with 1.2462, INTC with 1.3568 and CDR.WA with 1.0747. Also the Sharpe ratios of QCOM and ORCL are higher than the one of the aforementioned indices.
In summary the following conclusion can be drawn, from 16 investments exactly the half generated a Sharpe ratio above or nearly the same Sharpe ratio as a passive investment in indices, while the other half underperformed compared to the indices.

Nonetheless it is important to bench these Sharpe ratios also against other metrics. Cumming et al., (2021) investigated the Sharpe ratio for several hedge funds and indices for a shorter period, from 2008 until 2019, aligning more closely with the monitoring period used here. They found for the S&P 500 a annualized Sharpe ratio around $0.315 * \sqrt{12} = 1.09$ and for the Russell 2000 a Sharpe ratio of $0.221 * \sqrt{12} = 0.77$. They also reported a net-asset-value-weighted index, designed to mimic the hedge fund universe. For this Global Hedge Fund they reported a Sharpe ratio of $0.235 * \sqrt{12} = 0.81$. These are considerable higher Sharpe ratios, where none of the investments using the SADF test could surpass these aforementioned metrics. The CUSUM based investments still surpassing that threshold of the Global Hedge Fund five times and the high S&P 500 Sharpe ratio even two times. Therefore in more than half of the cases this approach delivered higher excess return per unit of risk than the average hedge fund depicted by the Global Hedge Fund in Cumming et al., (2021).

Next is a comparison with a simple buy and hold strategy. To make the returns comparable over different sampling periods the annual geometric average return is used. Compounding returns are computed since they better reflect the true return of the investment than using arithmetic average especially for volatile returns. Essentially the geometric average return is always less or at least as high as the arithmetic return and using the arithmetic average return would therefore bias the result upwards (Christopherson et al., 2009).
The results are given in table 8. All returns are calculated over the whole monitoring period.

Table 8

Returns

	SADF		CUSUM		Buy & Hold
Stock	**Geom. Avg.**	**Cumulative**	**Geom. Avg.**	**Cumulative**	**Total Return**
MSFT	12.58 %	231.60 %	24.84 %	839.66 %	1458.82 %
INTC	3.87 %	46.86 %	14.42 %	289.55 %	2282.01 %
QCOM	16.01 %	289.31 %	15.22 %	264.80 %	2785.71 %
ORCL	6.21 %	83.92 %	10.13 %	164.93 %	4419.63 %
CDR.WA	10.48 %	173.98 %	28.25 %	1132.90 %	1927.10 %
PLUG	18.00 %	184.02 %	2.44 %	16.39 %	631.95 %
VOW.DE	0.21 %	2.12 %	0.45 %	4.68 %	92.19 %
CGC	–0.48 %	–3.47 %	2.15 %	16.68 %	58.44 %

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–h; n.d.–i; n.d.–j; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details. Weekly price data is transformed into return series by the author. “Geom. Avg.” = annual geometric average return; “Cumulative” = cumulative return over the monitoring period. "Total Return" = return of holding the stock form the beginning until the end of the monitoring period. All values shown as percentages and rounded to two decimal places.

As stated in the beginning of the thesis the annualized geometric average return of the Russell 3000 index form January 1979 to December 2008 was 10.9% (Christopherson et al., 2009). It is evident from table 8 that using the SADF test the annual investment return surpassed 3 times (MSFT, QCOM and PLUG) the annualized return of the S&P 500 and provided one time with CDR.WA a nearly similar return to the index. However it yielded also 2 times very low returns with a near zero return for VOW.DE and a negative return for CGC of -0.48%. Just like in the Sharpe ratio measures the cumulative return and therefore also the annual geometric average return is at the CUSUM investments higher than the their SADF base counterparts for most stocks. Three CUSUM based investment yielded very low annual returns between 0.45% and 2.44%. Considering the bet from Warren Buffet also as benchmark, the highest compound annual return of the hedge funds was only 5.57% and of the index 7.1%. Compared to the hedge fund even 10 of the 16 investment surpassed that return. So only based on returns leaving risk aside the investments yielded very good returns overall. Especially the CUSUM yielded outstanding high returns in two cases with 28.25% annually for the CDR.WA investment and 24.84% for the MSFT investment.

Also included in table 8 are the returns by just holding the stock from the beginning of the monitoring until the end. Two things can be deduced from that. First, all the stocks which experienced a bubble still grow their stock prices over the whole monitoring period. Second the simple buy and hold strategies beat every of these investment strategies by a margin. The first implication shows kind of a selection bias as stocks are picked in this study which survived over a long period in markets which also grow themselves. Examples are ORCL, INTC or PLUG. To get a more complete picture on the returns of investments based on these tests also stocks without bubbles or stocks experiencing a bubble but dropping below their initial value need to be included. The second implication leads at first glance to the result that this trading strategy is not useful. However it needs to be said that against the other metrics above the investment delivered satisfactory results. The reason for that is probably that many investors professionals and beginners alike aren't holding stocks most of the time that long. There are several strong price movements in the "bubble phase" of the stocks. Especially these strong price movements in bubbles can lead to fear in investors and to a sudden sell of the stock not realizing the return depicted in table 8. Also on the other side there can be buy ins late in the upward movement of the bubble with a potentially high drop due to the collapse. This would be avoided by simply using the trading rule in this thesis and rule out the heuristic behavior seen in many instances by investors (Graham, 2016). Therefore these numbers should not seen as benchmark rather they are important for the first implication and so for further research.

5 Conclusion

In this thesis was investigated whether sequential bubble detection test can create a trading algorithm which generates excess returns when they are applied to stock data containing a bubble. To answer this question the SADF and CUSUM tests were applied each to historical data sets of eight stocks, using a real-time monitoring scenario. Everyone of these stocks experienced a sharp rise in their price at some point in the samples due to individual reasons.

To generate excess returns with this approach was partially achieved. The SADF based trading rule yielding only three stocks MSFT, QCOM and PLUG with Sharpe ratios in line with the Sharpe ratios of some indices but still lower than the Sharpe ratio of an average hedge fund. Despite that the CUSUM has delivered auspicious results. The Sharpe ratios of the investments in MSFT, INTC and CDR.WA were nearly as high or higher than the highest Sharpe ratio from the comparative data. In addition even the investments in 5 stocks adding to the aforementioned OCOM and ORCL surpassed the average hedge fund Sharpe ratio. Despite not being risk-adjusted and therefore violating the above definition of excess returns, also the annualized returns advocate the success of the CUSUM. Five of the eight investments delivered returns higher or nearly similar to what can be expected by The Russell 3000 between 1979 and 2008. Even 3 investments made by the SADF have surpassed the annual geometric average return of the Russell 3000.

But there is a shortcoming to answer the research question. First of all, the problem is that both tests lead to very different results regarding the geometric average returns or the Sharpe ratios. Such an investment approach can definitively yield higher returns than other investments and in some cases they are substantially higher. Like the investment in CDR.WA based on the CUSUM, which returned 28.25% per year and a Sharpe ratio of 1.3568. But the returns are definitely not consistent. For the CUSUM based investments PLUG, CGC and VOW.DE only realized a annual return between 0.45% and 2.44%, substantially less than any passive investing strategy in an index. The same can be said about the investments based on the SADF while here also one stock had a negative return, namely CGC with -0.48% annually. Additionally, all Sharpe ratios of the SADF were smaller than the Sharpe ratio of the average hedge fund depicted by the Global Hedge Fund. Also the CAPM results with overall only three abnormal returns out of the 16 investments are contradicting that the trading rule can yield excess returns. However as stated above the investment strategy had very low movements alongside the market and therefore these results don't seem utterly reflecting the investment performance.

One results which is not advocating that the strategy can yield excess returns but nevertheless is advocating sequential tests as a promising approach is that the SADF investment in the CGC stock is the only case with a negative return. In all other cases the strategy yielded positive returns indicating at least no substantial downward risk associated with this strategy when applied to stock with a bubble. A simple example shows why this is an important achievement. If 1000 euros are invested into an asset and this asset has decreased around 50% then just to get the initial 1000 euros back a 100% rebound of the stock is needed to get even again.

Consequently, these results shows that such sequential bubble detection tests can indeed return excess returns, but it definitely needs to be further studied to make them more consistent in delivering excess returns. Despite that, the study provided new insights into the practical application of this tests. It opens up a new strand of research that has been covered very sparsely, if not at all. Rather than repeating the same comparisons of which test is more reliable to detect a bubble in a historical data set. It has shown that these test can also lead to substantial returns by real-time monitoring stock prices with bubbles.

However there still needs to be extensive research regarding the application of these tests in real-time monitoring before some investors base their investments on such tests. First of all a comparison to other sequential tests is important. A comparison of other test to the SADF have bean made by Homm and Breitung (2012). Two test considered in their study are the Chow-Type Unit Root Statistic and the Busetti–Taylor statistic. They are also modified in a same manner the ADF is modified to become the SADF, by calculating the test statistic recursively and applying the supremum. They have higher finite sample power if the explosive behavior occurs late in the sample (Homm & Breitung, 2012). It would therefore be interesting for the further research if these other tests appear to be a more suitable choice for real-time monitoring applications, since speculative bubbles will certainly occur at the end of the sample in a real-time monitoring scenario. For the CUSUM test exists also an adjusted version which probably has higher power, the so called Backward Cumulative Sum (Otto & Breitung, 2022). This test leads to power gains compared to the conventional CUSUM, especially when the break is located at the end of a sample. Despite that the performance of the tests depends largely on a number of factors. The performance of others tests in this scenario is rather ambiguous simply for the reason that they are also used as an exit strategy.

Also to make this approach more suitable for investors it needs to be further investigated in different scenario. Further research needs to determine how well the investment based on such tests performs, applied to each stock of a whole index. One problem in this thesis is the selection bias mentioned before. By selecting only stocks which had a bubble and survived the whole monitoring period in the stock market, automatically many stocks selected had experienced a substantial upward shift in their prices. Obviously, this is intended since the thesis is a first attempt to research the return yielded by an investment in bubbles dictated by such tests. Nevertheless a normal investor will not know early enough whether a bubble occurs in a stock. Therefore by applying this to an index like the DAX this problem would be solved. A researcher would have stocks in the study which have their stock price plummeting during the monitoring period, resolving that problem of an general upward trend. Stocks which have a normal increase without a bubble and of course stocks containing a bubble. This would yield a more feasible approach for investors to investigate investment opportunities with these tests and would directly indicate whether the trading rule had outperformed passive investing in this index. Also the CAPM would be a better measure as the investments are all influenced by the same index. In such a scenario these tests would be very often applied and a good power and size becomes even more critical to not invest based on a false indication of explosive price patterns. Therefore it would be beneficial to connect this scenario with testing several structural break tests to identify clear differences in their reliability.

6 Reference Section

Astill, S., Harvey, D. I., Leybourne, S. J., Sollis, R. & Taylor, A. M. R. (2018). Real-Time Monitoring for Explosive Financial Bubbles. *Journal Of Time Series Analysis, 39*(6), 863–891. https://doi.org/10.1111/jtsa.12409

Astill, S., Harvey, D. I., Leybourne, S. J., Taylor, A. M. R. & Zu, Y. (2021). CUSUM-Based Monitoring for Explosive Episodes in Financial Data in the Presence of Time-Varying Volatility.*Journal Of Financial Econometrics, 21* (1), 187–227. https://doi.org/10.1093/jjfinec/nbab009

Aue, A. & Horváth, L. (2012). Structural breaks in time series. *Journal Of Time Series Analysis, 34*(1), 1–16. https://doi.org/10.1111/j.1467-9892.2012.00819.x

Berkshire Hathaway Inc. (2016). *Berkshire Hathaway Inc. 2016 Annual Report.* https://www.berkshirehathaway.com/2016ar/2016ar.pdf

Blanchard, O. & Watson, M. (1982). *Bubbles, Rational Expectations and Financial Markets.* https://doi.org/10.3386/w0945

Bodie, Z., Kane, A. & Marcus, A. J. (2005). *Investments*. McGraw-Hill US Higher Ed ISE.

Breitung, J. & Diegel, M. (2024), Sequential Detector Statistics for Speculative Bubbles, Available at SSRN: https://ssrn.com/abstract=5065510orhttp://dx.doi.org/10.2139/ssrn.5065510

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), 149–163. https://doi.org/10.1111/j.2517-6161.1975.tb01532.x

Campbell, J. Y., Lo, A. W. & MacKinlay, A. (1997). *The Econometrics of Financial Markets.* Princeton University Press. https://doi.org/10.2307/j.ctt7skm5

Chen, J. (2021, 4. Dezember). *Excess Returns Meaning, Risk, and Formulas*. Investopedia. https://www.investopedia.com/terms/e/excessreturn.asp

Christopherson, J. A., Carino, D. R. & Ferson, W. E. (2009). *Portfolio Performance Measurement and Benchmarking*. New York [u.a.]:Mcgraw-hill.

Chu, C. J., Stinchcombe, M. & White, H. (1996). *Monitoring structural change. Econometrica, 64*(5), 1045. https://doi.org/10.2307/2171955

Cumming, Douglas, Sofia Johan, and Geoffrey Wood (2021), *The Oxford Handbook of Hedge Funds*. https://doi.org/10.1093/oxfordhb/9780198840954.001.0001.

Duflo, E., Glennerster, R. & Kremer, M. (2006). Using Randomization in Development Economics Research: A Toolkit. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.951841

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review, 81*(4), 922–930. https://econpapers.repec.org/article/aeaaecrev/v_3a81_3ay_3a1991_3ai_3a4_3ap_3a922-30.htm

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25*(2), 383–417. https://doi.org/10.2307/2325486

Fama, E. F. & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *The Journal Of Economic Perspectives, 18*(3), 25–46. https://doi.org/10.1257/0895330042162430

Federal Reserve Bank of St. Louis. (n.d.). *Consumer Price Index for All Urban Consumers: All Items in U.S. City Average (CPIAUCSL)* [Historical data]. FRED, Federal Reserve Bank of St. Louis. Retrieved march 31, 2025, from https://fred.stlouisfed.org/series/CPIAUCSL

Grah, A. (2019, 23. Juli). Als Porsche einst Volkswagen übernehmen wollte. *dw.com*. https:

//www.dw.com/de/als-porsche-einst-volkswagen-%C3%BCbernehmen-wollte/a-49618351

Graham, B. (2016). *Intelligent investieren: Das Standardwerk des Value Investing* (9. Aufl.). FinanzBuch Verlag.

Harvey, D. I., Leybourne, S. J., Sollis, R.& Taylor, A. R. (2015). *Tests for explosive financial bubbles in the presence of non-stationary volatility. Journal Of Empirical Finance, 38*, 548–574. https://doi.org/10.1016/j.jempfin.2015.09.002

Hellrung, J. (2020, 3. April). CD Projekt: Warum der Publisher jetzt die wertvollste Firma Polens ist. *GameStar* Retrieved March 31, 2025. https://www.gamestar.de/artikel/cd-projekt-warum-die-cyberpunk-macher-ploetzlich-die-wertvollste-firma-polens-sind,3356223.html

Homm, U. and J. Breitung (2012). Testing for speculative bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Econometrics 10*(1), 198–231. https://doi.org/10.1093/jjfinec/nbr009

Hyndman, R.J., & Athanasopoulos, G. (2021) *Forecasting: principles and practice*, (3rd ed.), OTexts: Melbourne, Australia. Accessed on 04.04.2025 https://OTexts.com/fpp3

IBM. (n.d.). *Monte Carlo simulation* [Blog post]. IBM. Accessed on 14.04.2025 https://www.ibm.com/de-de/topics/monte-carlo-simulation

Jensen, M. C., Black, F. & Scholes, M. (1972). The capital Asset Pricing Model: Some Empirical tests. STUDIES IN THE THEORY OF CAPITAL MARKETS, *Praeger Publishers Inc*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=908569

Kindleberger, C. P. & Aliber. (2005). *Manias, Panics and Crashes.* (5th ed.) [eBooks]. Palgrave Macmillan. https://doi.org/10.1057/9780230628045

Kirchgäßner, G., Wolters, J. Hassler, U. (2007). Introduction to modern Time Series analysis. In *Springer eBooks*. `https://doi.org/10.1007/978-3-540-73291-4`

Kortian, T. (1995). RDP 9501: Modern Approaches to Asset Price Formation: A Survey of Recent Theoretical literature. *RePEc: Research Papers in Economics.* `https://www.rba.gov.au/publications/rdp/1995/pdf/rdp9501.pdf`

Kurozumi, E. (2020). Asymptotic properties of bubble monitoring tests. *Econometric Reviews, 39*(5), 510–538. `https://doi.org/10.1080/07474938.2019.1697086`

MacDonald, C. (2024, Dezember 12). Could the Cannabis Bubble Re-Inflate? *The Motley Fool Canada.* Retrieved March 31, 2025 `https://www.fool.ca/2024/12/12/could-the-cannabis-bubble-re-inflate/`

Meyer, S., Uhr, C., Johanning, L. (2021). *Private investors and the emergence of neo-brokers: Does payment for order flow harm private investors?* `https://assets.traderepublic.com/assets/files/202111_study_private_investors_and_the_emergence_of_neo_brokers.pdf`

Page, E. S. (1954). Continuous inspection schemes. *Biometrika, 41*(1/2), 100. `https://doi.org/10.2307/2333009`

Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika, 42*(3–4), 523–527. `https://doi.org/10.1093/biomet/42.3-4.523`

Pavlidis, E., Martínez-García, E. & Grossman, V. A. (2018). Detecting periods of exuberance: A look at the role of aggregation with an application to house prices. *Federal Reserve Bank Of Dallas, Globalization And Monetary Policy Institute Working Papers, 2017*(325). `https://doi.org/10.24149/gwp325r1`

Phillips, P., Wu, Y. and Yu, J. (2011). EXPLOSIVE BEHAVIOR IN THE 1990s NASDAQ: WHEN DID EXUBERANCE ESCALATE ASSET VALUES?*. *International Economic Re-*

view, 52(1), 201–226. https://doi.org/10.1111/j.1468-2354.2010.00625.x

Phillips, P., Shi, S. & Yu, J. (2015a). 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

Phillips, P. C. B., Shi, S. & Yu, J. (2015b). TESTING FOR MULTIPLE BUBBLES: LIMIT THEORY OF REAL-TIME DETECTORS. *International Economic Review, 56*(4), 1079–1134. https://doi.org/10.1111/iere.12131

Otto, S. & Breitung, J. (2022). BACKWARD CUSUM FOR TESTING AND MONITORING STRUCTURAL CHANGE WITH AN APPLICATION TO COVID-19 PANDEMIC DATA.*Econometric Theory, 39(4)* , 659–692. https://doi.org/10.1017/s0266466622000159

Stock, J. H. & Watson, M. W. (2014). *Introduction to Econometrics* (3rd ed.). Pearson International.

Tsay, R. S. (2005). Analysis of Financial Time Series (2nd ed.). *Wiley series in probability and statistics.* https://doi.org/10.1002/0471746193

Vasilopoulos, K., Pavlidis, E. & Martínez-García, E. (2022). exuber: Recursive Right-Tailed Unit Root Testing with R. *Journal Of Statistical Software, 103*(10). https://doi.org/10.18637/jss.v103.i10

Yahoo Finance. (n.d.-a). *NASDAQ Composite (^IXIC))* [Historical data]. Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/^IXIC/history/

Yahoo Finance. (n.d.-b). *S&P 500 (^GSPC)* [Historical data]. Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/^GSPC/history/

Yahoo Finance. (n.d.-c). *CBOE Interest Rate 10 Year T No (^TNX)* [Historical data]. Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/^TNX/history/

Yahoo Finance. (n.d.-d). *Microsoft Corporation (MSFT)* [Historical data].Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/MSFT/history/

Yahoo Finance. (n.d.-e). *Intel Corporation (INTC)* [Historical data].Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/INTC/history/

Yahoo Finance. (n.d.-f). *QUALCOMM Incorporated (QCOM)* [Historical data]. Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/QCOM/history/

Yahoo Finance. (n.d.-g). *Oracle Corporation (ORCL)* [Historical data].Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/ORCL/history/

Yahoo Finance. (n.d.-h). *CD Projekt S.A. (CDR.WA)* [Historical data]. Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/CDR.WA/history/

Yahoo Finance. (n.d.-i). *Plug Power Inc. (PLUG)* [Historical data].Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/PLUG/history/

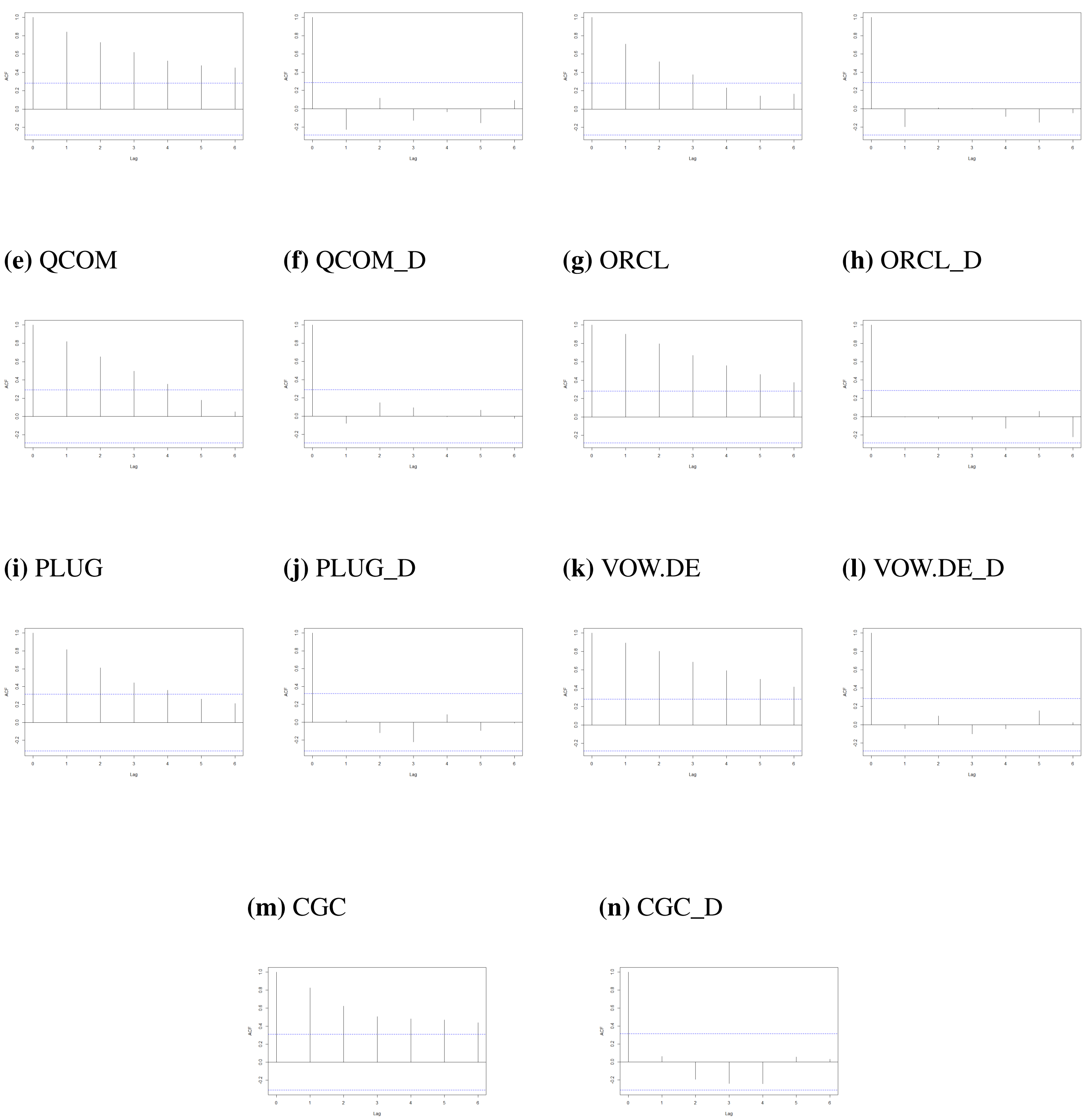
Yahoo Finance. (n.d.-j). *Volkswagen AG (VOW.DE), Historical data*.Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/VOW.DE/history/

Yahoo Finance. (n.d.-k). *Canopy Growth Corporation (CGC)* [Historical data]. Yahoo. Retrieved march 31, 2025, from https://finance.yahoo.com/quote/CGC/history/

7 Appendix

Figure 9

ACF of Original and Differenced Price Series



(a) MSFT **(b)** MSFT_D **(c)** INTC **(d)** INTC_D

(e) QCOM **(f)** QCOM_D **(g)** ORCL **(h)** ORCL_D

(i) PLUG **(j)** PLUG_D **(k)** VOW.DE **(l)** VOW.DE_D

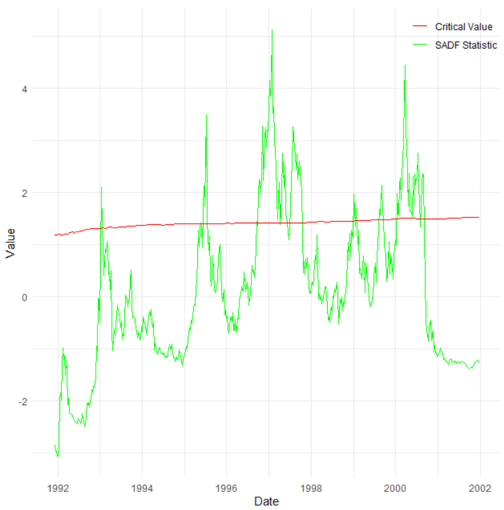
(m) CGC **(n)** CGC_D

Note. Data from Yahoo Finance (n.d.–d; n.d.–e; n.d.–f; n.d.–g; n.d.–i; n.d.–j; n.d.–k). Retrieved March 31, 2025. See Table 2 for full details. The ACF plots are calculated by the author based on weekly price data. Plots for original series are labeled with the acronym, and plots for differenced series are labeled with the acronym followed by "_D".

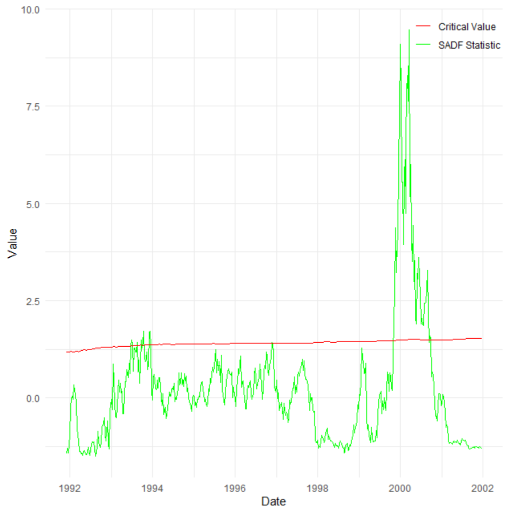
Figure 10

Critical Values vs. Test Statistic (SADF)

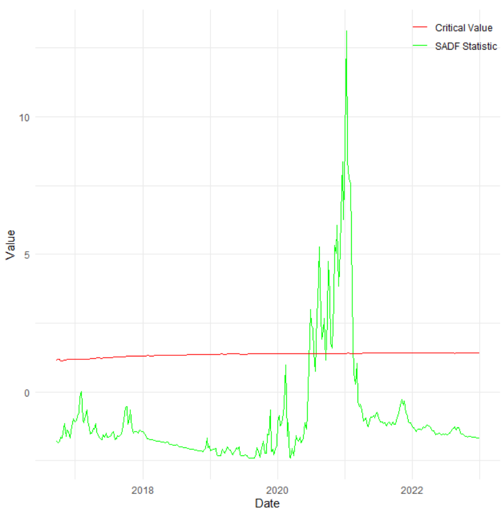
(a) INTC



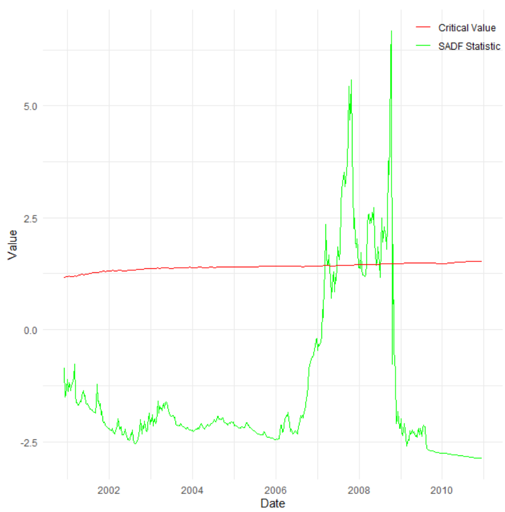
(b) ORCL



(c) PLUG



(d) VOW.DE

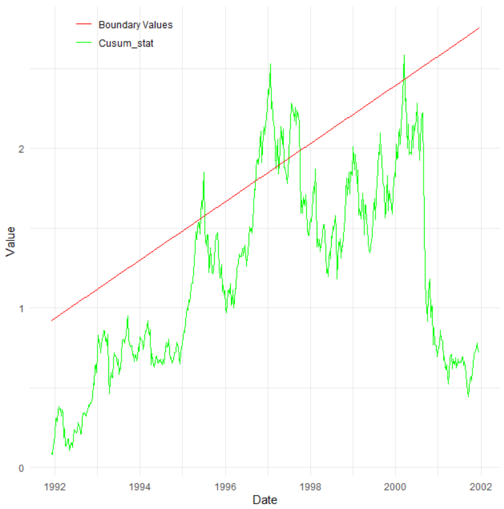


Note. Data from Yahoo Finance (n.d.–e; n.d.–g; n.d.–i; n.d.–j). Retrieved March 31, 2025. See Table 2 for full details. SADF test statistics is calculated directly from weekly raw price data. Depicted are here the 95% finite sample critical values calculated trough Monte Carlo simulations. Both calculated by the author.

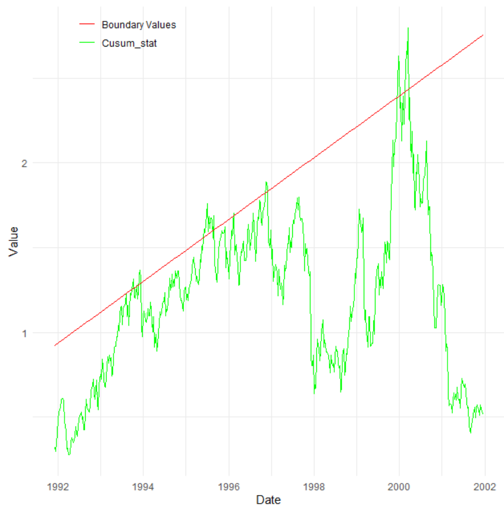
Figure 11

Critical Values vs. Test Statistic (CUSUM)

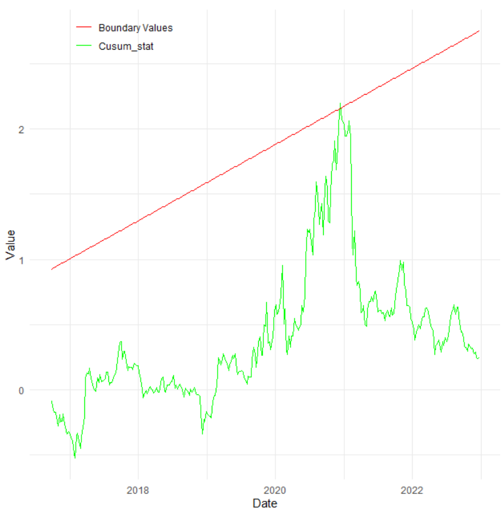
(a) INTC



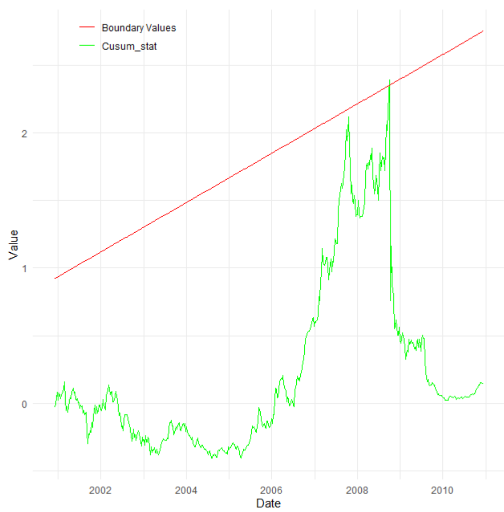
(b) ORCL



(c) PLUG



(d) VOW.DE



*Note.*Data from Yahoo Finance (n.d.–e; n.d.–g; n.d.–i; n.d.–j). Retrieved March 31, 2025. See Table 2 for full details. CUSUM test statistics were calculated by the author directly from weekly raw price data and compared against the derived critical boundary function with the corresponding 95% critical value for γ_{aplha}.