On Explosive Time Series

by

Ioannis Korkos

This thesis is submitted for the degree of
Doctor of Philosophy in Finance

Essex Business School
University of Essex

June 2020

Abstract

The first chapter of this thesis, discusses the characteristics of an asset bubble episode outlining the reasons these episodes have attracted so much interest nowadays and provides an overview of historical bubble episodes motivating the testing procedures proposed in Chapters 2-4.

The second chapter proposes a right-tailed bootstrap implementation of the covariate Augmented Dickey-Fuller (CADF) unit root test of Hansen (1995), motivated by the work of Chang, Sickles and Song (2017). We apply the right-tailed bootstrap BCADF test in a recursive manner and provide evidence that the inclusion of relevant covariates offers significant power gains. An empirical application of the proposed methodology is conducted, utilising the Moody's Seasoned Aaa and Baa Corporate Bond Yields, the Ten-Year Treasury Rate and the Volatility Index (VXO) as covariates.

The third chapter intends to examine the size and power properties of right-tailed Dickey-Fuller unit root test processes when testing for market efficiency in the commodity markets by applying a wild bootstrap approach to Phillips et al. (2015) tests. The simulations results show that the proposed wild bootstrap test offers better size control and power performance in finite samples. In the empirical exercise, our proposed test suggests periods of market inefficiency prior to the existence of the bubble episode as identified by the conventional tests during two periods of oil crises.

The fourth chapter studies the hypothesis of an asset bubble in a rational expectations framework using a bivariate coexplosive vector autoregression as in Nielsen (2010). Firstly, we apply a co-explosive vector autoregression to model whether the WTI crude oil price run-up of 2007-2008 can be attributed to the existence of a bubble as well as whether the WTI crude oil collapse of 2014-2015 exhibits characteristics of bubble implosion.

In the fifth and final chapter, concluding remarks are made regarding and directions for future research are proposed.

Acknowledgements

Throughout the writing of this thesis I have received a great deal of support and assistance without which this thesis would be incomplete. First and foremost, I am deeply indebted to my PhD advisers Professor Neil Kellard and Dr Sam Astill for their continued support, help, inspiration, guidance and friendship during the past three years.

Furthermore, I am very grateful to Professor A. M. Robert Taylor for providing constructive and useful feedback on chapters 2 and 3 of this thesis. I would like to thank Dr Daniel Tsvetanov as well for his support in the earlier stage of this thesis. Earlier versions of the chapters of this thesis has been presented to international conferences where great feedback and suggestions for improvements have been taken. In particular, special thanks should be given to Professor Peter C. B. Phillips, Professor Liudas Giraitis and Professor Roderick McCrorie for their very insightful remarks and comments. Additionally, I would like to express my gratitude to Professor Bruce E. Hansen, Professor Wonho Song and Professor A. M. Robert Taylor, for kindly sharing their GAUSS codes with me. Moreover, I would like to gratefully acknowledge the financial support of the Economic and Research Council (award reference 1797607), during the writing of this thesis.

Moreover, I am thankful to both the PhD viva external examiner, Professor Anastasios Magdalinos and the internal examiner Professor Ekaterini Panopoulou for their useful feedback and comments as well as Dr Theodora Bermpei for conducting the mock viva.

Finally and most importantly, I would like to express my sincere gratitude to my family for their love and unconditional support during this academic journey. I am grateful to Pori Saikia, Lizeta Demetriou, Maris Camilleri and Lucas Shuttleworth for their emotional support, love and kindness without which this thesis might not have become complete and my PhD colleagues at the Essex Business School for all the good times and long days we shared together.

Declaration

Chapter 2 of this thesis is based on the paper "Improving the power of univariate tests for bubbles by using covariates" co-authored by myself, Sam Astill, Neil Kellard and A. M. Robert Taylor. Chapter 3 of this thesis is based on the paper "Wild Bootstrap Testing for Speculative Bubbles Using Spot and Futures prices" co-authored by myself, Sam Astill and Neil Kellard. Chapter 4 of this thesis is based on the paper "Testing for Bubbles in Commodity Spot and Futures Using a Co-explosive Autoregression" co-authored by myself, Sam Astill and Neil Kellard. In all instances I was the primary author of the paper.

This work is, to the best of my knowledge original, except were acknowledgements and references are made to previous work. Neither this, nor any substantially similar thesis has been or being submitted for any other degree, diploma or other qualification at any other university.

Contents

1 Introduction **10**

1.1 Bootstrap Unit Root Testing for Explosive Behaviour using Covariates . 14

1.2 Wild Bootstrap Testing for Speculative Bubbles Using Spot and Futures Prices 16

1.3 Testing for Bubbles in Commodity Spot and Futures Using a Co-explosive Autoregression 18

1.4 Contribution to the extant literature 22

2 Bootstrap Unit Root Testing for Explosive Behaviour Using Covariates 23

2.1 Introduction 23

2.1.1 Theory 23

2.1.2 Early Tests on Bubbles 23

2.1.3 Recent Tests on Bubbles 26

2.2 The Model and Assumptions 35

2.3 Extant Recursive Test Procedures 37

2.4 Extant Covariate Augmented Unit Root Tests 39

2.5 Proposed Tests 39

2.5.1 Limit Theory 40

2.5.2 Practical Implementation of tests 41

2.5.3 Bootstrap Unit Root Tests with Covariates 42

2.5.4 Bootstrap Limiting Distribution 45

2.6 Finite Sample Simulations 45

2.6.1 Empirical Size 47

2.6.2 Empirical Power 48

2.7 Empirical Application 49

2.8 Conclusion 53

2.9 Tables 55

2.10 Figures 59

3 Wild Bootstrap Testing for Speculative Bubbles Using Spot and Futures Prices **62**

3.1 Introduction 62

3.1.1 Rational Bubble Tests 62

3.2 The Model and Assumptions 66

3.3 Extant Recursive Test Procedures 68

3.4 Limit Theory 69

3.5 Proposed Tests 70

3.5.1 Wild Bootstrap Unit Root Tests 71

3.6 Finite Sample Simulations 74

3.6.1 Empirical Size 75

3.6.2 Empirical Power 77

3.7 Empirical Application 78

3.7.1 Testing for Autocorrelation 82

3.8 Conclusion 83

3.9 Tables 85

3.10 Figures 89

4 Testing for Bubbles in Commodity Spot and Futures Using a Co-explosive Autoregression 109

4.1 Introduction 109

4.1.1 Cointegration in a Rational Expectations Framework 109

4.1.2 Analysis of Coexplosive and Cointegrated Processes 110

4.1.3 Recent Unit Root Tests for Explosive Behaviour 112

4.2 The Model and Assumptions 116

4.2.1 The VAR Model 117

4.2.2 The Cointegrated VAR Model 117

4.2.3 The Coexplosive VAR Model 117

4.3 The Granger-Johansen Representation 119

4.4 Statistical Analysis and Hypothesis Testing 120

4.5 Limit Theory 122

4.6 Empirical Application 123

4.6.1 Data 123

4.6.2 Cointegration Tests 123

4.6.3 Coexplosiveness: the contemporaneous case 124

4.6.4 Coexplosiveness: the non-contemporaneous case 125

4.6.5 Causality Testing 128

4.6.6 Unit Root Testing for Rational Bubbles 130

4.7 Conclusion 134

4.8 Tables 137

4.9 Figures 146

5 Concluding Remarks 155

5.1 Future Research 158

A Appendix The Limit Theory of the CGSADF and the bootstrap CGSADF test 160

List of Tables

1	Finite Sample Size	56
2	Full Sample Unit Root Test	57
3	Bubble Date Stamping	58
4	Finite Sample Size	86
5	Bubble Date Stamping (weekly frequency)	87
6	Bubble Date Stamping (monthly frequency)	88
7	Cointegration rank determination for the period July 2007 to July 2008, contemporaneous series	138
8	Cointegration rank determination for the period July 2007 to July 2008, non-contemporaneous series (lags)	139
9	Cointegration rank determination for the period November 2015 to February 2016, contemporaneous series	140
10	Cointegration rank determination for the period November 2015 to February 2016, non-contemporaneous series (lags)	141
11	Characteristic Roots	142
12	Granger Causality test	143
13	Unit Root Tests	144
14	Bubble Explosion Date Stamping	145
15	Bubble Implosion Date Stamping	145

List of Figures

1	Finite Sample Power	60
2	Finite Sample Power	61
3	Finite Sample Power	90
4	Finite Sample Power	91
5	Finite Sample Power	92
6	Bubble Date Stamping (weekly frequency)	93
7	Bubble Date Stamping (weekly frequency)	94
8	Bubble Date Stamping (weekly frequency)	95
9	Bubble Date Stamping (weekly frequency)	96
10	Bubble Date Stamping (weekly frequency)	97
11	Bubble Date Stamping (monthly frequency)	98
12	Bubble Date Stamping (monthly frequency)	99
13	Bubble Date Stamping (monthly frequency)	100
14	Bubble Date Stamping (monthly frequency)	101
15	Bubble Date Stamping (monthly frequency)	102
16	Sample Autocorrelation Function (weekly frequency)	103
17	Sample Autocorrelation Function (weekly frequency)	104
18	Sample Autocorrelation Function (weekly frequency)	105
19	Sample Autocorrelation Function (monthly frequency)	106
20	Sample Autocorrelation Function (monthly frequency)	107
21	Sample Autocorrelation Function (monthly frequency)	108
22	WTI crude oil prices	147
23	Johansen Cointegration Test on WTI crude oil prices	148
24	Bubble Explosion Date Stamping	149
25	Bubble Explosion Date Stamping	150
26	Bubble Explosion Date Stamping	151
27	Bubble Implosion Date Stamping	152
28	Bubble Implosion Date Stamping	153
29	Bubble Implosion Date Stamping	154

1 Introduction

Asset price bubbles have recently attracted significant interest in the finance literature as their collapse has a significant impact on the real economy. According to economic theory, an asset bubble can be defined as a prolonged period of substantial price deviations from a fundamental value (see *inter alia* Blanchard and Watson 1982, Campbell, Lo and MacKinlay 1997 and Homm and Breitung 2012). Price misalignments lead investors to pay a higher price (than justified by fundamentals) for an asset, expecting to sell the asset at an even greater price in the future and generate a profit. In a bubble regime, there is a high volume of trading, in contrast to normal market conditions (Ofek and Richardson 2003). Subsequently, positive feedback mechanisms result in further inflation of the equilibrium price.

Historical episodes of price bubbles have been well documented in literature, see *inter alia* Galbraith (1997), Kindleberger and Aliber (2005), and Sornette (2003b). The earliest known bubble episode in the financial history is known as Tulip Mania and took place in the Netherlands during the 17th century. In the 18th century, the first significant market crash in the British stock market occurred, driven by what is known as the South Sea Bubble; a result of excessive speculation by the South Sea Company that had monopolistic rights to shipping and trading activities with South America (Sornette 2003b). A similar bubble episode occurred in France over the exact same period. In this case, banks excessively issued bank notes which were not equivalent to their gold and silver reserves. In 1720, the market crashed and this event is known historically as the Mississippi Bubble (Kindleberger and Aliber 2005).

Moving into the twentieth century, during the Roaring 20s the U.S. economy was thriving mainly as a result of new technological innovations and industrialisation. In 1929, the Federal Reserve of the U.S. attempted to calm down the market through implementing tight monetary policy. Panic resulted in a massive liquidation of shares, margin investors went bankrupt and major banks were driven into default since they had invested depositors' money, leading to a recession that lasted almost four years, known as Great Depression. At the beginning of the Great Moderation there is the market crash of October 1987, referred to as Black Monday, came after a period of euphoria in capital markets as a result of low interest rates, mergers and acquisitions, hostile takeovers and leverage buyouts. The Federal Reserve of the U.S. increased interest rates and made access to funding extremely unaffordable. Computer trading (sell orders after losses), derivative securities, liquidity problems, huge trade and budget deficits and overvaluation combined with austere monetary policy led two of the largest capitalisation indexes in the U.S. (S&P 500 and Dow Jones) to a decline of more than 20% of their value (Sornette 2003b).

The term "dot-com" bubble or "tech" bubble is widely used to describe the last few years of the 1990s, a decade when the stock prices of internet firms escalated to extremely high levels. In a short period of time, hundreds of thousands of small-medium sized firms raised capital through IPOs despite cash flow issues, taking advantage of the enthusiasm of capital markets participants to fund internet firms. During the early 2000s, investors realised that the price of many internet stocks was well above their fundamental value, with the price of these stocks subsequently crashing, resulting in a mild recession for the U.S economy, despite the effort of the Federal Reserve of the U.S. to decrease interest rates (Ofek and Richardson 2003, and Kindleberger and Aliber 2005).

More recently, the global financial crisis of 2007-2008 has been triggered by the sub-prime mortgage market crash (Akerlof and Shiller 2009). Due to the deep integration of the capital and money markets nowadays, the exuberance was transmitted from the financial markets (commodities, exchange rates, stock exchanges etc.) to the real economy. At the end of 2006, the mortgage backed security market reached extremely high levels of volume and at the same time, the majority of debtholders were unable to pay-back their debt leading to delinquencies and foreclosures. Investors lost trust, liquidity sank and the financial system, especially investment and commercial banks, collapsed. The contagion propagated to commodities, real exchange, fixed income and oil markets as investors selectively transferred their assets to other investments. This global financial crisis drove the majority of the developed countries into recession (Phillips & Yu, 2011). Akerlof and Shiller (2009) attribute the recent global financial crash to the breakdown of the financial system and especially of structured financial products, the high leverage and capital loss of the financial institutions and the already-agreed credit lines between the banks and their clients.

The main focus this thesis is on identifying explosive episodes in financial time series. Our approach is twofold. At first, we concentrate on improving the size and power performance of the Phillips et al. (2015) tests by including covariates (Chapter 2) and by applying a wild bootstrap procedure (Chapter 3) to the standard Phillips et al. (2015) tests. Then, we emphasize on long-term relationships between assets and we study the existence of an explosive root in a cointegrating relationship between two financial series (Chapter 4).

In particular, putting emphasis on the early detection of asset price bubbles in Chapter 2, we investigate whether the power of right-tailed Dickey-Fuller unit root test procedures can be improved by the inclusion of relevant covariates. Choosing to test for a bubble in a univariate framework can lead to potential power reductions since ignoring any correlation with other time series could possibly have a negative impact on the explanatory power of standard unit root tests leading to significant power losses (Hansen, 1995). Applying sub-sample techniques may result in imprecise estimation of the nuisance parameter introducing additional variability and causing severe size distortions (Chang, Sickles and Song 2017). Dealing with a nuisance dependency problem, we apply a bootstrap procedure ensuring the asymptotic validity of the critical values from the distribution of the test statistics that lead to improved size and power performance in finite samples. In our empirical exercise, we manage to detect earlier two major explosive episodes: the Black Monday of October 1987 and the dot-com bubble.

On the same framework, examining the size and power properties of right-tailed Dickey-Fuller unit root test processes when testing for market efficiency in the commodity markets in Chapter 3, we apply a wild bootstrap approach to Phillips et al. (2015) tests to account for potential heteroskedasticity that resembles the pattern of structural breaks, regime changes or volatility breaks offering robust critical values. In fact, the wild bootstrap test appears to control for size better than the non-bootstrap test while the power performance is significantly improved as we model the series of interest as a moving average process rather than a unit root process since under the null hypothesis of market efficiency the expected future spot price should equal the price of the futures contract. Applying the test empirically, we identify the 2007-2008 oil price run-up and the 2014-2015 oil price collapse while the conventional test of Phillips et al. (2015) does either not identify any episode at all or identifies the episode with delay, reflecting the superior power of our proposed wild bootstrap test to effectively identify such episodes.

Following a VAR approach in Chapter 4, we allow for explosive roots as suggested by

Nielsen (2010) while testing for cointegration offering the advantage of performing the cointegration analysis of Johansen (1991) even in the presence of explosive behaviour in the related series. Looking into the WTI crude oil market, we find that both oil prices of spot and futures contracts are $I(1, x)$ processes and the two variables cointegrate such that their linear combination is an $I(0)$ process for the period July 2007-July 2008. Our empirical findings are in accordance to Pavlidis et al. (2017) since there is no statistical evidence of explosive behaviour on the differences between the future spot price and the futures contract price for that period and therefore the linear relationship is stationary.[1]

Chapter 5 concludes and discusses some avenues for future research.

[1]The notation $I(1, x)$ stands for variables with both explosive and random walk components and $I(x)$ for variables with just explosive common trends as in Nielsen (2010).

1.1 Bootstrap Unit Root Testing for Explosive Behaviour using Covariates

The identification of asset price bubbles is clearly of great interest for both theorists and empirical researchers. Recently, many econometric techniques have been developed for bubble detection in the context of time series analysis. Cointegration analysis, for instance, has been applied as one of the main testing approaches assessing price deviations from equilibrium (see *inter alia* Campbell and Shiller 1987, Campbell and Shiller 1988a and Campbell and Shiller 1988b). In a rational bubble regime, the equilibrium condition between the asset price and its market fundamental is violated. Therefore, non-stationary deviations from the general equilibrium in the long term provide evidence in favour of a bubble. Non-stationary behaviour is examined in the logarithmic transformation of the price dividend ratio through unit root testing as well; if the dividend yield is integrated of order one, then this could be considered as a strong evidence of a rational bubble (Campbell and Shiller 1987). The rational expectations theory perceives (rational) bubbles as anticipated phenomena where their expected value of next period depends on the compounded value of the bubble at time zero. In other words, the value of a bubble today equals the discounted value of future bubble episodes.

With the need for early detection of asset price bubbles apparent, this chapter intends to investigate whether the power of right-tailed Dickey-Fuller unit root test procedures can be enhanced by the inclusion of relevant covariates. If choosing to test for a bubble in a univariate framework, examining a variable in isolation can be rather costly in terms of power since ignoring any correlation with other time series could possibly weaken the explanatory power of standard unit root tests leading to significant power losses (Hansen, 1995).

Our proposed Covariate Augmented Dickey Fuller unit root test is applied in a backward supremum sequence as suggested by Phillips et al. (2015) to detect explosive episodes that occur at the end of the sample. Applying sub-sample techniques can, however, lead to imprecise estimation of the nuisance parameter introducing additional variability and causing severe size distortions (Chang, Sickles and Song 2017). To deal with the nuisance dependency problem we apply a bootstrap procedure, ensuring the asymptotic validity of the critical values drawn from the bootstrap distribution of the test statistics. We concentrate on the case where the explosive episode takes place at the end of the sample to date-stamp bubbles in real time. The simulations show that the proposed bootstrap tests offer impressive size and power performance in finite samples. In particular, the bootstrap tests appear to be less size distorted compared to the non-bootstrap conventional unit root tests and the inclusion of covariates in the standard Augmented Dickey Fuller regression model offers significant power gains.

We conduct empirical work to investigate the effectiveness of the proposed tests on the early identification of bubble episodes. Specifically, we examine whether our proposed tests would have detected known past bubbles in the S&P 500 price dividend series before the tests of Phillips et al. (2015) when used as an early warning mechanism, utilising the Moody's Seasoned Aaa and Baa Corporate Bond Yields, the Ten-Year Treasury Rate and the Volatility Index (VXO) as covariates. The superiority of our proposed test is reflected on the earlier detection of two major explosive episodes: Black Monday of October 1987 and the dot-com bubble.

1.2 Wild Bootstrap Testing for Speculative Bubbles Using Spot and Futures Prices

A large number of studies has recently focused on studying asset bubbles as their collapses can have significant impact on the real economy whereas the identification and dating of the bubble episodes is of particular importance for investors, policy makers and central banks. One aspect of asset bubbles that is particularly interesting for academics and researchers are asset bubbles that hold in a rational expectations framework. In this framework, real asset prices should be equal to the present value of the future cash flows, the fundamentals, that the asset generates, augmented by a bubble component that grows with the real interest rate in the presence of rational bubbles.

Interestingly, Diba and Grossman (1988) tested for the presence of rational bubbles in stock prices, suggesting that persistent explosive behaviour that cannot be differenced to stationary might indicate the existence of rational bubbles. Their approach has been subject to criticism by Evans (1991) on the grounds of the poor power performance of traditional unit root tests to identify explosive episodes that collapse periodically in the sample. More recently, literature has concentrated on applying right-tailed unit root tests to the level of a series with Phillips et al. (2011) introducing a forward recursive right-tailed supremum ADF test that has good power properties and is fairly simple to use. Since then, the weight of interest has been shifted to identifying bubble episodes on a real time basis rather than historical episodes. Phillips et al. (2015) propose a backward recursive right-tailed supremum ADF test that is rather useful on date-stamping past bubble episodes and a generalised double-recursive right-tailed supremum ADF test that has better size and power performance in identifying multiple bubble episodes in the sample.

Stressing out the importance of early identification of asset bubble episodes, this chapter intends to examine the size and power properties of right-tailed Dickey-Fuller unit root test processes when testing for market efficiency in the commodity markets by applying a wild bootstrap approach to Phillips et al. (2015) tests to account for potential heteroskedasticity that might be attributed to structural breaks, regime changes or volatility breaks as the wild bootstrap procedure can resemble the behaviour of a time series that has heteroskedastic innovation terms while offering robust critical values. For this reason, we model the series of interest as a moving average process rather than a unit root process since under the null hypothesis of market efficiency the difference between the future spot price and the futures contract price will be a moving average process of order determined by the length of the futures contract.

We concentrate on the case where the explosive episode occurs at the end of the sample to identify these episodes in real time. The simulations results show that the proposed wild bootstrap test offers better size control and power performance in finite samples since the standard backward recursive right-tailed supremum ADF test considers that the series is a unit root process under the null hypothesis when the series is in fact stationary, whereas the wild bootstrap implementation of the backward recursive right-tailed supremum ADF test simulates critical values under the null hypothesis that the series is a moving average process with the order of that process depending on the length of the futures contract.

Particularly, the wild bootstrap test appears to be less size distorted compared to the non-bootstrap test while the power gains are significantly higher. In the empirical exercise, testing for market efficiency in the commodity markets we apply the proposed and extant tests on the difference between the WTI crude oil future price and the price of nine futures contracts across different maturities over the period September 1995 to July 2019. Focusing mainly on the 2007-2008 oil price run-up and the 2014-2015 oil price collapse, our proposed test identifies the two episodes while the conventional test of Phillips et al. (2015) does either not identify an episode at all, or identify the origination day of the episode with delay reflecting the superior power of our proposed wild bootstrap test to effectively identify episodes of non-stationarity that occur at the end of the sample. The proposed test suggests periods of market inefficiency prior to the existence of the bubble episode as identified by the conventional tests.

1.3 Testing for Bubbles in Commodity Spot and Futures Using a Co-explosive Autoregression

Recent unprecedented imbalances in the financial markets have attracted significant interest from professionals, regulators and a growing number of academics as they might exhibit asset bubble characteristics. In contrast to the consensus that in time-series econometrics, variables are either stationary or second order integrated, speculative bubbles in prices result in an explosive root in addition to a unit root making testing for cointegration and statistical inference rather inconclusive.

In the econometrics literature, it is commonly argued that the existence of speculative bubbles imply no cointegration between prices and fundamentals (see *inter alia* Diba and Grossman 1988b). In contrast, in the presence of speculative bubbles, prices and fundamentals can be cointegrated so their linear combination does not contain a unit root while at the time there is an explosive root in the system (Nielsen 2010).

In a cointegration framework, both the explosive and unit root need to be tested. Nielsen (2010) suggests that the cointegrated vector autoregression introduced by Johansen (1991) can be used in a context where some of the series are integrated of order greater than one. Therefore, even though one of the series might be explosive, the Johansen (1991) cointegrated VAR model can still estimate the cointegrating relationship given, of course, that the two series are cointegrated. Nielsen (2010) introduces the idea of coexplosiveness to allow the standard cointegrated VAR models to test for the existence of bubbles. In particular, Nielsen (2010) proposes a VAR model that allows both unit roots and explosive characteristic roots, utilising the standard cointegration techniques introduced by Johansen (1991). The coexplosive and cointegrated vector autoregressive model arises as a restriction to the standard VAR model and allows both a random walk and an explosive stochastic component with a characteristic root larger than one. In other words, the cointegrated VAR approach of Johansen (1991) offers the advantage of testing for a unit root between two series and simultaneously testing for an explosive root in at least the one of the two series (Nielsen 2010).

This approach contradicts Diba and Grossman (1988) on the fact that two series can be cointegrated and yet, their linear combination might contain an explosive component (Engsted 2006). As a result, the VAR approach developed by Johansen (1991) offers the advantage of testing for stationarity while simultaneously testing whether at least one of the variables has an explosive characteristic root since testing for the number of cointegrating vectors in the coexplosive case is similar to the standard Johansen (1991) procedure. The reason for this is that the asymptotic distribution of the likelihood ratio test when there is an explosive root is the same as in the standard Johansen cointegration test (Nielsen 2010).

In this chapter we utilise the application of Nielsen (2010) approach to test for cointegrating relationships across different series while simultaneously testing whether the series contain any explosive components, allowing to perform the cointegration analysis of Johansen (1991) even in the presence of explosive behaviour in the related series. Particularly, we utilise Johansen's cointegration rank test to analyse the oil price run-up in the WTI crude oil market between July 2007 and July 2008 as this period is indicated as explosive as well as the oil price collapse between November 2015 and February 2016, contributing to the debate of whether the 2007-2008 oil price run-up can be attributed to the existence of a speculative bubble as well as whether the oil price collapse of 2014-2015 exhibits any characteristics of bubble implosion.

We find that in the contemporaneous case crude oil spot prices and all futures contracts contain both an explosive root and a unit root component from July 2007 to July 2008, whereas when we match the futures contract prices with the actual future spot prices then oil future prices of spot and the prices of the six month, twelve month and eighteen month futures contracts contain both an explosive root and a unit root component during this period.

Concerning the 2014-2015 crude oil price collapse we argue that contemporaneously, crude oil spot prices and the one month futures contract and crude oil spot prices and the three month futures contract contain both an explosive root and a unit root component between November 2015 and February 2016, whereas matching the futures contract prices with the actual future spot prices results in a single explosive root between the future spot prices and the three month futures contract therefore the system contains both an explosive root and a unit root component for this period.

Our empirical findings suggest that both oil prices of spot and futures contracts are $I(1,x)$ processes and the two variables cointegrate such that their linear combination is an $I(0)$ process for the periods July 2007 to July 2008 and November 2015 to February 2016 for some of the futures contracts, in support of the view commonly stated in the empirical literature that prices of spot and (short maturity) futures contracts should be cointegrated even when there is a bubble episode in the sample (Engsted 2006).

Investigating the oil price run-up of 2007-2008 and oil price collapse of 2014-2015 further, we extend our analysis to study Pavlidis et al. (2018), according to which it is the fundamentals that are responsible for the oil price run-up during the early 2000s and not the existence a speculative bubble. In particular, we apply the ADF, SADF and GSADF tests to the difference between the future spot prices and futures contract prices across all different maturity contracts for the sample period September 1995 to July 2019. Our cointegration analysis seems to be in accordance to Pavlidis et al. (2017) since applying the BSADF test when the test is applied on the difference between the future spot prices and the futures contract prices provides no statistical evidence of explosive behaviour between July 2007 and July 2008 and November 2015 to February 2016 that we identify coexplosiveness, implying that the linear relationship is stationary, although date-stamping only identifies the origination date of the bubble episode with delay across futures contract with different maturities.

Applying a date-stamping technique to the difference between the future spot prices and the futures contract prices results in a delayed identification of the origination date of the bubble oil episode of 2007-2008 providing no statistical evidence of explosive behaviour between July 2007 and July 2008. Furthermore, applying the same date-stamping technique to the reverse series of the difference between the future spot prices and the futures contract prices results in a delayed identification of the origination date of the oil price collapse episode of 2014-2015 providing no statistical evidence of explosive behaviour (in the reverse series, therefore no market collapse in the original series as in Phillips and Shi, 2018) between November 2015 and February 2016. These findings are in support of our evidence that during the peak of the oil price run-up of 2007-2008 and the oil price collapse of 2014-2015, crude oil future spot prices and futures contract prices are cointegrated, therefore their linear relationship is stationary and since the characteristic roots of their VAR model are, in some cases, explosive we conclude that oil prices of the spot and futures contract coexplode during these two periods.[2]

The results of this chapter seem to be in line with the standard present value model

[2] Note that coexplosiveness in the reverse series means that the two series collapse together so that their cointegrating relationship still holds, they co-implode. During the period November 2015 and February 2016 we find coexplosiveness in the reverse series of the difference between the future spot prices and the futures contract prices and therefore co-implosiveness in the original (non-reversed) series.

as augmented by a bubble component to account for a rational bubble. The linear combination of spot and futures contract prices contains an explosive root as a result of a speculative bubble. Examining variables in a bivariate framework might offer significant advantages as bubble episodes emerging in the futures market might be transmitted in the spot market causing speculative bubbles and thus cointegration and coexplosive analysis can be proven very valuable in bubble identification.

1.4 Contribution to the extant literature

This thesis contributes to the extant literature in various ways. Firstly, the novelty of the second chapter is the suggestion of a multivariate approach to existent tests for bubble identification by introducing covariates in a recursive framework. We deal with, potential, imprecise estimation of the nuisance parameter that causes inaccurate statistical inference and severe size distortions due to sub-sample testing by applying a bootstrap technique ensuring the asymptotic validity of the critical values drawn from the bootstrap distribution. We improve the size and power performance of the existent tests while empirically we identify two historical episodes, namely the Black Monday of October 1987 and the dot-com bubble earlier compared to the bubble detection tests suggested in the econometrics literature.

In the third chapter, we consider a wild bootstrap approach to existent tests for bubble identification to account for the possibility of heteroskedastic residuals that can be attributed to breaks in volatility in order to study market efficiency in the commodity markets. The simulation results suggest that the wild bootstrap test offers improved size control while offering significant power gains as the series of interest has been modelled as a moving average process rather than a unit root process since under the null hypothesis of market efficiency the difference between the future spot price and the futures contract price will be a moving average process of order specified by the length of the futures contract. In the empirical application, our proposed wild bootstrap test identifies periods of market inefficiency prior to a bubble episode that existent bubble detection tests do not detect, acting as an early warning mechanism of non-stationary behaviour in the market that could, potentially, lead to a bubble episode.

In the forth chapter, we examine questions of bubble identification and market efficiency using a bivariate approach, in contrast to the extant literature that studies asset price bubbles in a univariate framework. Firstly, we apply a co-explosive vector autoregression to test whether the WTI crude oil price run-up of 2007-2008 and the oil price of collapse of 2014-2015 can be attributed to the existence of a bubble. We find that there is an explosive root in the system and that oil spot and futures contract prices at various maturities, are cointegrated over that period. Secondly, we apply recent univariate bubble tests to test for market efficiency. We conclude with an evaluation regarding the most appropriate approach to bubble identification in commodity markets.

2 Bootstrap Unit Root Testing for Explosive Behaviour Using Covariates

2.1 Introduction

Financial bubbles have recently attracted a considerable amount of research work in both the economics and financial literature. A financial bubble is commonly defined as a sudden, continuous rise of the price of one or multiple assets. Speculators' activity triggers further rises of the price that results in a crash due to reversals of expectations (Kindleberger 1987).

2.1.1 Theory

Under the assumption of a rational bubble regime, the real price of an asset equals the present value of its relevant fundamentals (Lucas 1978) e.g. dividends (i.e. expected value of future cash flows).[3] In their seminal work, Campbell and Shiller (1987) have meticulously studied the validity of this present value model assuming either a constant discount rate and implying that the two series are cointegrated[4], under the assumption that the transversality condition[5] holds, or a time-varying one, arguing that in this case the logarithmic difference between prices and dividends is stationary.

The persistent failure of the present value models to justify deviations from fundamentals that lead to bubbles resulted in the development of methods for detecting explosive episodes that mainly focus on the rational bubble assumption.

2.1.2 Early Tests on Bubbles

In that framework, Shiller (1981) suggests a methodology that takes into account the variance bounds of stock prices to evaluate the present value model. However, this approach can only provide point estimates of variance and therefore hypothesis testing cannot be utilised. For this reason, LeRoy and Porter (1981) generate estimates of variances for stock prices and dividends in a bivariate framework. Although this method

[3]Deviations from equilibrium due to non-fundamental determinants can as well be integrated into the standard present value model by dropping the transversality condition.

[4]Therefore drifts away from fundamentals are corrected in the long-run.

[5]The transversality condition provides a unique solution of the present value model, thus the equity price equals the market fundamental price whereas in case that the condition does not hold a set of solutions are given.

was not initially designed for bubble testing, useful implications can be derived when in a bubble regime, the variance of the asset price exceeds the variance justified by fundamentals. However, the variance bound tests as proposed by Shiller (1981) are subject to criticism; Flood and Hodrick (1986) argue that these particular tests are not appropriate for bubble testing and might mislead by providing evidence for bubble existence due to misspecification errors and (or) incorrect modelling of expectations.

First to introduce bubbles in the alternative hypothesis, West (1987) presents a two-step test which requires the specification of an equilibrium model and investigates the impact of the fundamental value on the asset price, given the Euler equation as a no-arbitrage asset pricing model. The difference between the two estimates (the actual and the constructed one) of the effect of dividends on the asset price can be attributed to either model misspecification or bubble. Flood et al. (1994) criticise this approach emphasising that even after performing the misspecification tests, rejections might be justified by other factors such as the inadequacy of the rational models to explain bubbles or the invalidity of the standard asymptotic inference resulting from the non-ergodic data generation processes.

From what stated above, it is clear that conventional univariate econometric techniques suffer from a series of problems such as omitted variable biases which might lead to the rejection of the null hypothesis of no bubble (Flood and Garber 1980), model misspecifications or inconsistent statistical tests (Flood and Hodrick 1986), low power on identifying rational bubbles, especially in a periodically collapsing framework (Evans 1991), and (or) size distortions and low power (see *inter alia* Stock, 1991, Campbell and Perron 1991, Domowitz and El-Gamal 2001).

In their seminal research, Diba and Grossman (1988) highlight the significance of unit root testing into rational bubble detection. They introduce a standard left-tailed unit root process to test for the null hypothesis of a random walk, under the assumption of a time-invariant discount rate to argue that there is no evidence of bubble existence if both stock prices and dividends are non-stationary in levels but stationary in differences. Furthermore, Diba and Grossman (1988) apply standard unit root tests to the real S&P 500 stock price index between 1871 and 1986, finding that stock prices and dividends are non-stationary in levels but stationary in differences, concluding that a rational asset bubble can be identified when a time series cannot be differentiated to stationarity, due to the explosiveness of the dataset. In the context of longer-term relationships between two variables, Diba and Grossman (1988) utilise the Bharghava (1986) ratios[6] to conclude that evidence of cointegration indicates no bubble, in disagreement with Evans (1991) who argues that traditional unit root tests are non-capable of capturing complex bubble characteristics due to the non-linear structure of the bubble models.

Hall et al. (1999) suggest a generalised version of the ADF test which incorporates the dynamic Markov regime-switching models as proposed by Hamilton (1989) and Hamilton (1990). This approach is consistent with the argument that an explosive autoregressive root indicates the existence of rational bubbles. Furthermore, Van Norden and Vigfusson (1998) argue that size distortion has a considerable impact on regime-switching models and suggest the utilisation of the Van Norden (1996) test instead, which assess the switching probabilities depending on the size of the bubble. The benefits of introducing stochastic regime-switching models into the bubble identification process are pointed out by Driffill and Sola (1998) who advocate that deviations of stock prices from fundamentals can be perceived as shifts in fundamentals due to regime change and not as bubble phenomena. However, there is evidence that Markov-switching models might infer false detection or spurious explosiveness (Shi 2013).

Several research work has focused on cointegration, long memory and persistence. Campbell and Shiller (1987) introduce the idea of cointegration between stock prices and dividends as an evidence of no bubble. In the case of a rational bubble, the long-term relationship between prices and fundamentals (e.g dividends) is violated and prices move away from equilibrium for a prolonged period of time. Mixed results in their empirical part lead Campbell and Shiller (1987) to conclude that the drifts away from fundamentals are quite persistent, although highly sensitive to the discount rate.

In a fractional integration framework, Cuñado et al. (2005) follow a fractionally integrated methodology in stock prices and dividends of NASDAQ to conclude that the sampling frequency of the data affects the statistical inference of bubble existence. In particular, testing by using daily and weekly data provides evidence of fractional cointegration, since the order of integration lies between zero and one whereas testing at monthly frequency results in no rejection of the unit root hypothesis of no cointegration. The authors attribute this distortion to either bias resulting from the usage of

[6] von Neumann-type ratios as invariant hypotheses testing processes.

low-frequency data, known as the temporal aggregation problem or sample size. This bias might lead to wrong inference of slow convergence or random walk (Taylor, 2001). Persistent trend-cyclical behaviour seems to fade out when the same data are examined for longer periods (Mandelbrot, 1969).

In the same framework, Koustas and Serletis (2005) apply fractional integration techniques in the logarithmic dividend yield of the S&P 500 finding evidence of long memory against rejecting the null hypothesis of a rational bubble. On the contrary, Frömmel and Kruse (2012) criticise the methodology proposed by Koustas and Serletis (2005) arguing that possible structural breaks are not taken into consideration and they suggest a test for changing persistence under fractional integration based on Sibbertsen and Kruse (2009) accounting for both long memory and changing persistence, combining structural breaks and unit root testing in accordance to Demetrescu et al. (2008). Gürkaynak (2008) emphasises that the degree of integration of the unobservable fundamentals can be greater than one, explaining the inference of non-stationarity.

Moving from rational bubbles, Froot and Obstfeld (1991) introduce the concept of intrinsic bubbles, defined as episodes of exuberance caused by exogenous economic fundamentals to describe nonlinear fluctuations in asset prices. They empirically test the proposed model in the US stock exchange market and attribute the existence of bubbles to the nonlinearities between equity and stock prices, arguing that the proposed tests utilise estimates consistent under both the null and the alternative hypothesis.

Extending Campbell and Shiller's (1987) cointegration restriction by imposing a robust no rational bubble constraint which does require neither a constant discount factor nor a specific asset pricing model, Craine (1993) provides evidence that the discount factor for the S&P 500 can be non-stationary and therefore any inference of bubble might be misleading.

2.1.3 Recent Tests on Bubbles

Traditional unit root tests may lead to spurious indications of explosive behaviour in the presence of non-stationary volatility. Cavaliere and Taylor (2008) suggest a new set of approaches to unit root testing which deal with permanent volatility shifts. Particularly, rather than performing the ADF tests directly on the original time series, they implement these tests on the inverted time transformation of the original time series resulting in good power gain. Furthermore, Cavaliere and Taylor (2009), investigate the case of a near-unit-root process under the assumption of non-stationary autoregressive volatility. In particular, under weak dependence they suggest a wild bootstrap method and provide empirical evidence in accordance to the proposed methodology which performs good under the presence of near-integrated autoregressive stochastic volatility with leverage effects.

The power properties of the unit root tests are of particular interest. Leybourne (1995) introduces a joint test which employs both a reverse and a forward ADF unit root test, taking advantage of the fact that under the null hypothesis both tests are marginally asymptotically distributed in an identical way, however there is no perfect correlation due to the different ending points of the time series examined (Leybourne & Taylor 2003). The proposed methodology of Leybourne (1995) chooses the maximum value of the forward and reverse ADF test-statistics and rejects the null hypothesis more often compared to the standard ADF test. Furthermore, Leybourne (1995) compares this joint max ADF test with the standard ADF test to demonstrate that it offers significant power gains with similar size properties. Forward and reverse estimation is of particular usefulness in seasonal unit root testing as well. Leybourne and Taylor (2003) introduce a combination of the Hylleberg et al. (1990) seasonal unit root testing (HEGY) for both the forward and reverse processes. Testing on the power and finite-size properties of the new model through Monte Carlo simulations, they infer that the inclusion of the Leybourne (1995) joint test into seasonal unit root testing offers superior size and power gains compared to other OLS or weighted symmetric least squares (WSLS) processes.

Examining the asymptotic distribution of a random walk, Abadir and Lucas (2000) use the unit root M-tests to derive the limit theory that depends on a nuisance parameter. Furthermore, they argue that in the random walk case the limit distribution does not follow a standard normal distribution but a skewed one instead, highlighting a nuisance dependency problem and they derive a normal approximation for the quantiles of the test-statistics that are based on robust unit root M-estimators. Following Elliott (1998), Magdalinos and Phillips (2009) highlight the problem of asymptotic bias of integrated, near-integrated or explosive regressors and provide limit theory, extending Phillips and Magdalinos (2008) theory for cointegrated systems that are fully explosive. The relationship between the explosive regressors defines the asymptotic behaviour of the least squares estimator of the cointegrating coefficients. Furthermore, Magdalinos and Phillips (2009) show that in the moderately explosive case, the OLS regression process is asymptotically median unbiased and the limit theory is mixed normal. Finally, they highlight that in a moderately explosive framework the regressors might appear to be explosively cointegrated due to contagion effects on other variables.

Recent research focuses on right-tailed unit root tests, which have the property of detecting mildly explosive or sub-martingale behaviour in time series by putting emphasis on the alternative hypothesis of explosiveness. Phillips et al. (2011) introduce the argument that explosiveness in asset prices and not in fundamentals may indicate a bubble episode, suggesting a forward recursive right-tailed supremum Augmented Dickey-Fuller (SADF) test, capable of detecting ongoing bubbles. Phillips et al. (2011) apply this test on the NASDAQ stock price and dividend index, finding evidence of the dot-com bubble in the early 2000s. The main advantage of the Phillips et al. (2011) methodology is that it can be successfully utilised not only in stock prices but in commodity future prices (Gilbert 2010), commodity and house prices (Homm and Breitung 2012), the exchange rate market (Bettendorf and Chen 2013) and crude oil spot and future markets (Tsvetanov et al. 2016) as well.

Homm and Breitung (2012) perform simulations and investigate the power properties of a Chow-type Dickey-Fuller test, a modified version of the locally best invariant (LBI) test of Busetti and Taylor (2004) and the Phillips et al. (2011) test to conclude that the Phillips et al. (2011) methodology can be a powerful tool not only in a structural break environment but on identifying end-of-sample bubble episodes as well. Investigating the validity of the tests empirically, they find strong evidence of explosiveness in the pre-2008 subprime mortgage downturn in the UK, US and Spanish house markets, in accordance to their main argument that bubble episodes occurred in multiple markets.

In a multiple bubble environment however, bubble detection can be challenging and the Phillips et al. (2011) unit root process may be less successful and powerful on identifying multiple bubble phenomena. The conventional Augmented Dickey-Fuller unit root test may infer pseudo stationarity; evidence of stationary behaviour when the data is non-stationary. In order to overcome these weaknesses of the Phillips et al. (2011) methodology in a multiple bubble regime, an extension of the Supremum ADF test is introduced. Phillips et al. (2015) construct the Phillips et al. (2015) test, a consistent technique of identifying multiple bubbles with periodically collapsing behaviour not only in historical prices but on a real-time basis as well. In particular, they introduce two recursive window processes; a generalised and a backward version of the SADF test (GSADF and BSADF respectively).

According to Harvey et al. (2017) the power of the two tests mentioned above, depends on the location of the explosive regime and on whether there is a collapse inside the sample as well. In multiple bubble and collapses, the GSADF test outperforms, in terms of size and power, the conventional (Phillips et al. 2011) SADF test by recursively changing the starting and ending points of the sample covering more subsamples of the data. Furthermore the GSADF test is designed to detect the existence of one or more explosive episodes in a financial time series that can occur anywhere in the sample. In addition, the GSADF can only show whether there is a bubble episode in the sample without indicating the location the episode occurs within the sample. The BSADF test is a backward recursive right-tailed ADF test and it is a more powerful detection tool of bubble episodes that occur at the end of the sample and has been developed by Phillips et al. (2015) for date-stamping of the origination and termination dates of the explosive episode.

Phillips et al. (2015) test empirically the validity of their proposed models to the S&P 500 stock price and dividend index between January 1871 and December 2010 and find evidence of explosive behaviour by applying the GSADF test. Then they utilise the BSADF test for samples ending in each time period in order to date-stamp these events and they successfully detect more than six historical banking crises and bubble episodes in this time spam. However, there seems to be a delay bias in the detection of the explosive episodes in Phillips et al. (2015); Phillips and Shi (2018) argue that there is a delay on estimating the dates of crisis origination and market recovery dates and thus suggest a methodology based on "reverse regression" strategies. Focusing on various different ways of bubble implosion that mainly depend on the mature of the collapse and trying to deal with bias, Phillips and Shi (2018) incorporate a reverse sample-regression into the recursive window process of Phillips et al. (2015). At the same time they embody a market recovery parameter which is the date that asset prices return to equilibrium, on the Phillips et al. (2015) methodology and following Rosser (2000) and Huang et al. (2010), Phillips and Shi (2018) distinguish market crashes into "sudden", "disturbing" and "smooth" reflecting different ways of price decline. Furthermore, they define market recovery as the date when the asset prices of a particular market return to their "normal martingale path" or alternatively to fundamentals (equilibrium). The innovation of this approach is that it differentiates the date of the bubble implosion from the market recovery date, defining the latter as the ending point of the mildly unit root collapsing process. The reverse-regression strategy can offer valuable information in a multiple bubble framework where the number of explosive episodes and (or) collapses are not known in advance.

A main advantage of this methodology compared to the Harvey et al. (2012) and Harvey et al. (2015) test is that it is capable of estimating the amount of bubble episodes and crash points at the same time whereas the Harvey et al. (2012) and Harvey et al. (2015) approach requires the number of these episodes to be known beforehand. Testing their model on the NASDAQ stock market index for the period January 1973 to August 2013, Phillips and Shi (2018) uncover four different stages of the dot-com episode of explosive behaviour: the origination date (December 1996), the implosion date (February 2000), the market correction or recovery date (December 2000) and finally a further correction date (February to April 2004). The utilisation of the reverse-regression implementation strategy on the right-tailed unit root testing can offer significant information on market recovery due to the high sensitivity of the right-tailed testing processes to deviations from equilibrium. Therefore, the reverse regression procedure can be considered as a real-time technical analysis of explosive episodes in financial markets.

Moreover, Astill et al. (2017) account for conditional and unconditional heteroskedasticity and serial correlation on end-of-sample explosive episodes of financial time series by introducing Andrews (2003) and Andrews and Kim (2006) end-of-sample instability tests. Their proposed methodology can offer significant power gains compared to the BSADF test of Phillips et al. (2015) and can be utilised as an early-detection mechanism of end-of-sample bubble episodes.

As important as it is to detect periods of explosiveness, it is crucial as well to be able to assess the origination and termination of a bubble regime precisely. Harvey et al. (2012) introduce an alternative approach in right-tailed unit root testing by using the minimum sum of squared residuals estimators (see *inter alia* Bai and Perron 1998 and Kejriwal, Perron, and Zhou 2013) together with the Bayesian Information Criterion (BIC) for the optimum lag length selection. The proposed methodology can perform well in detecting bubbles that collapse within the sample as well as end-of-sample bubble episodes.

Phillips (2016) relates asset bubbles with the sentiments of heterogeneous investors who perceive fundamentals in different ways and therefore overreact (exuberant), underreact (cautious) or respond appropriately to changes in fundamentals (fundamentalists), whereas Lee and Phillips (2016) advocate that myopic investing introduces speculative behaviour into asset pricing and provide a finite investor horizon study. Although the validity of the standard present value model in the long term may be considered as given, the low power of the unit root tests, possible non-linear relationships, structural breaks and possible outliers may lead to rather ambiguous or mistaken inferences (Bohl and Sicklos 2004).

From what has been stated so far, we may argue the right-tailed unit root tests can be considered as quite powerful bubble detection mechanisms that succeed to detect not only past bubbles but bubble episodes that grow in real time as well. Nevertheless, in a univariate framework the convention to ignore any correlation with other time series may be costly; the exclusion of a correlated stationary covariate from the standard regression model could weaken the explanatory power of the unit root testing which would lead to significant power losses (Hansen, 1995). Examining a variable in isolation is a rather simplistic approach in time series analysis and the inclusion of a highly correlated stationary variable could offer significant reduction in error variance compared to the standard ADF test.

In addition, Hansen (1995) argues that with the existence of covariates, the critical values of the standard ADF test can lead to incorrect statistical inference. Additionally, as discussed later, the form of the asymptotic distribution of the conventional ADF test provides conservative asymptotic critical values and therefore the ADF test appears to have low power. Therefore, Hansen (1995) suggests using the first differences of a covariate before including it into the regression model in order to deal with non-stationarity and derives the asymptotic distribution of the covariate Augmented Dickey-Fuller (CADF) test, which is a convex combination of the standard Dickey-Fuller distribution and the normal distribution. By using a Monte Carlo simulation process in a no-deterministic environment, Hansen (1995) concludes that the CADF is more powerful compared to the standard ADF test.

Although, it is widely known that the unit root testing processes suffer from low power, the ADF test has the lowest size-adjusted power amongst the unit root tests, asymptotically and in finite samples, it provides the least size distortions as well (Stock 1994).[7] Caporale and Pittis (1999) advocate that Hansen's (1995) CADF test succeeds in achieving both great power gains and small size distortions, in contrast to the conventional unit root tests, resulting in less over-rejection of the alternative hypothesis of stationarity. Moreover, they provide theoretical evidence of reduction of the standard errors and coefficient estimates that depends on the contemporaneous and temporal correlation structure of the errors and the stationary covariate. Furthermore, Caporale and Pittis (1999) argue that the value of the test-statistic of the CADF test will converge to that of the standard univariate ADF test where there is neither contemporaneous nor temporal correlation between the covariate and the error term. In addition, they apply both the ADF and the CADF tests to the Nelson and Plosser (1982) dataset[8] and conclude that the inclusion of covariates might not only enhance the explanatory power of the model but reverse the presumption of unit root as well. This is in accordance to their argument that by using the CADF test, the power of the model can be improved with relatively small size distortions.

In a vector autoregressive framework, Elliott and Jansson (2003) extend the CADF test proposed by Hansen (1995) to account for the case where constants or time trends are included in the unit root regression model. In addition they suggest a likelihood-ratio-based approach in combination with a GLS demeaning/detrending process for the dependent variable and an OLS demeaning/detrending process for the covariate. Westerlund (2013) follows a similar GLS process combined with an ARCH to account for heteroskedasticity.

Chang, Sickles and Song (2017) bring the literature's attention to an important caveat; the limit distribution of the covariate ADF test depends on the correlation between the equation error and the covariate. Therefore, the nuisance parameter de-

[7] Stock (1994) develops a comparative study of univariate unit root tests with similar local asymptotic power functions but different finite-sample behaviour to infer that the DF test-statistic exhibits the least size distortions (compared to other unit root tests), at the expense of low power.

[8] Nelson and Plosser (1982) perform the conventional ADF test to fourteen US macroeconomic time series (e.g. GDP, employment) to infer that there is evidence of non-stationarity in all of them except unemployment.

pendency results in invalid statistical inference and large size distortions. In order to ensure the asymptotic validity of the critical values, they suggest a parametric bootstrap CADF test which improves the explanatory power of the model (especially in the case where the covariate is highly correlated with the error term), with no effect on the size properties of the standard ADF test.[9] Moreover, Chang, Sickles and Song (2017) apply this bootstrap method to an extension of the Nelson and Plosser (1982) dataset and argue that the bootstrap CADF test can offer statistical gains compared to the sample CADF test due to the independence of bootstrapped critical values from the nuisance parameter. The extended Nelson-Plosser (1982) dataset was firstly introduced by Schotman and Van Dijk (1991) and it is widely used in macroeconomic analysis. The data series includes a variety of macroeconomic variables -the nominal and real GNP, employment, industrial production and money stock among others- as covariates. Additionally, Aristidou, Harvey and Leybourne (2017) consider a GLS-demean/detrend and an OLS-demean/detrend CADF approach in the existence of asymptotically non-negligible initial conditions in order to obtain efficient estimates and improve the power of the CADF test, as proposed by Elliott, Rothenberg and Stock (1996), which depends on the local asymptotic power and the magnitude of the initial conditions.

Nowadays, the weight has partially shifted from rational to irrational bubbles with behavioural finance attributing bubble detection to behavioural determinants such as mimetic and herding aspects of the investors' attitude (see *inter alia* Akerlof and Shiller 2009, Shiller 2015). Avery and Zemsky (1998) advocate that asset bubbles and excess volatility can be partially explained by herd behaviour. In particular, they examine whether herding can act as a triggering factor of a bubble crash and in contrast to the general conviction, they conclude that models of rational trading can indeed interpret herding and crashes. Abreu and Brunnermeier (2003) relate price bubbles to the asynchronous selling strategies of rational arbitrageurs as well as the different beliefs on the timing of the bubble burst.

Overall we can argue that the time-varying present value model can provide valuable information for the stock price behaviour in the long term (see *inter alia* Campbell and Shiller 1988 and Campbell and Shiller 1989) when in the short term, deviations from

[9] Chang, Sickles and Song (2017) consider the time trend case as well and demonstrate that the inclusion of trend in the unit root regression equation results in power loss for all tests except the bootstrapped CADF test.

fundamentals can be attributed to non-fundamental determinants; speculative bubbles (West, 1987 and Evans, 1991), noise trading (Shleifer 2000), expansions and recessions in business activity (Phelps and Zoega 2001) or deviations that might lead to crashes (Charemza and Deadman 1995). The theoretical framework of behavioural characteristics such as overconfidence, enthusiasm, greed, fear and panic can trigger the development of financial bubbles and deviations from general equilibrium, has been developed by Akerlof and Shiller (2009).

In this chapter we examine whether the power and size performance of right-tailed Dickey-Fuller unit root test procedures can be enhanced by the inclusion of relevant covariates. As argued by Hansen (1995), in a univariate framework examining a variable in isolation can lead to power reduction since ignoring information in correlated series could possibly weaken the explanatory power of the standard unit root tests. Furthermore, we apply the proposed Covariate Augmented Dickey Fuller unit root test in a backward supremum sequence as in Phillips et al. (2015) to identify explosive episodes that occur at the end of the sample. Sub-sample techniques can lead to imprecise estimation of the nuisance parameter introducing additional variability, causing severe size distortions as discussed in Chang, Sickles and Song (2017).

As a remedy to the nuisance dependency problem we apply a bootstrap procedure while ensuring the asymptotic validity of the critical values drawn from the bootstrap distribution of the test statistics. We put emphasis on the case where the explosive episode takes place at the end of the sample as date-stamping bubbles in real time can be of great usefulness to policy makers and central banks. The simulations suggest that the proposed bootstrap tests offer great size and power performance in finite samples as the bootstrap tests appear to be less size distorted compared to the non-bootstrap conventional unit root tests and therefore the inclusion of covariates in the standard Augmented Dickey Fuller regression model offers significant power gains.

In our empirical application, we investigate whether our proposed tests would have detected known past bubbles in the S&P 500 price dividend series before the tests of Phillips et al. (2015) when used as an early warning mechanism, utilising the the Moody's Seasoned Aaa and Baa Corporate Bond Yields, the Ten-Year Treasury Rate and the Volatility Index (VXO) as covariates. The superiority of our proposed test is reflected on the earlier detection of Black Monday of October 1987 and the dot-com bubble.

In Section 2.1 we discuss some theory together with earlier and more recent developments in the bubble identification literature. In Section 2.2 we outline the explosive financial bubble model and in Sections 2.3 and 2.4 we review the extant covariate and recursive unit root tests respectively. In Section 2.5 we present our proposed tests, some limit theory and our bootstrap approach, whereas in Section 2.6 the finite and sample size and power properties are examined by using Monte Carlo simulations and the relevant discussion takes place. Section 2.7 presents an empirical application of the proposed tests. Section 2.8 summarizes and concludes. Tables and Figures are presented in sections 2.9 and 2.10 respectively. Mathematical proofs are given in the Appendix.

In what follows $\xrightarrow{p}$ denotes convergence in probability, $\xrightarrow{d}$ denotes convergence in distribution and $\lfloor . \rfloor$ denotes the integer part of its argument. We denote $\|.\|$ as the Euclidean norm as well. For a vector $z = z_i$, $\|z\|^2 := (\sum_i z_i^2)^{1/2}$ and for a matrix $A = (a_{ij})$ $\|A\|^2 := (\sum_{i,j} a_{i,j}^2)^{1/2}$.

2.2 The Model and Assumptions

Consider a time series process $\{y_t\}$ that consists of a purely deterministic component and a stochastic component generated according to the following data generating process (DGP);

$$y_t = d_t + S_t, \; t = 1, ..., T, \tag{2.1}$$

where d_t is the deterministic component and can be either equal to 0 (neither constant, nor trend), μ (constant but no trend) or $\mu + \theta t$ (constant and trend). The initial condition y_0 is assumed to be stochastically bounded and does not affect the subsequent analysis in this paper. The stochastic component, $\{S_t\}$, is generated according to;

$$\Delta S_t = \delta S_{t-1} + u_t. \tag{2.2}$$

The innovation sequence, $\{u_t\}$ is generated according to

$$\alpha(L) u_t = b(L)' \Delta x_t + \varepsilon_t \tag{2.3}$$

where Δx_t is an m-vector of stationary covariates, $\alpha(L)$ is a lag operator polynomial of order p: $\alpha(z) = 1 - \sum_{k=1}^{p} \alpha_k z_p^k$ and $b(k) = \sum_{k=-r}^{q} \beta_k z^k$ is a polynomial allowing for, but not requiring, both leads and lags of Δx_t to enter the DGP.

In the context of testing the null hypothesis of a unit root against the alternative of stationarity Hansen (1995) combines (2.2) and (2.3) and proposes estimating the following regression by OLS

$$\Delta y_t = d_t^\dagger + \delta y_{t-1} + \sum_{k=1}^{p} \alpha_k \Delta y_{t-k} + \sum_{k=-r}^{q} \beta_k' \Delta x_{t-k} + \varepsilon_t =: CADF(p, r, q). \tag{2.4}$$

where

$$d_t^\dagger = \begin{cases} 0 & \text{if } d_t = 0 \\ -\delta\mu & \text{if } d_t = \mu \\ a(1)\theta - \delta\mu - \delta\theta & \text{if } d_t = \mu + \theta t \end{cases} \tag{2.5}$$

Following Chang, Sickles and Song (2017) we assume that the stationary covariates Δx_t are generated by an $AR(\ell)$ process given by

$$\Psi(L)\Delta x_{t+r+1} = \eta_t, \tag{2.6}$$

where $\Psi(z) = I_m - \sum_{k=1}^{\ell} \Psi_k z^k$.

We also make the following assumptions on the innovation sequence $\xi_t = (\varepsilon_t, \eta_t')'$ that defines the correlation between the stationary covariates Δx_t and the series of interest $\{y_t\}$.

Assumption 2.1. *(a) Let $\{\xi_t\}$ be a martingale difference sequence such that $E(\xi_t\xi_t') = \Sigma$ and $(1/T)\sum_{t=1}^{T} \xi_t\xi_t' \xrightarrow{p} \Sigma$ with $\Sigma > 0$ and $E|\xi_t|^\gamma < K$ for some $\gamma > 4$, where K is some constant depending only upon γ*

(b) $\alpha(z)$, $det(\Phi(z)) \neq 0$ for all $|z| \leq 1$

Remark 2.1. *As noted by Chang, Sickles and Song (2017) Assumption 2.1(a) allows for conditional heteroskedasticity, including GARCH behaviour, in all equations in the system including the covariates. By definition (ε_t) is uncorrelated with (η_{t+k}) for $k \geq 1$. This condition implies that (ε_t) is uncorrelated with the lagged differences* *of the dependent variable* $(\Delta y_{t-1}, ..., \Delta y_{t-p})$ *and the leads and lags of the covariates* $(\Delta x_{t+r}, ..., \Delta x_{t-q})$.

Remark 2.2. *As noted by Chang, Sickles and Song (2017) Assumption 2.1(a) implies the following invariance principle*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor rT \rfloor} \xi_t \xrightarrow{d} B(r) \tag{2.7}$$

holds for $r \in [0.1]$ as $T \to \infty$. The process $B(r) = (B_\varepsilon, B'_\eta)'$ is an $(1+m)$-dimensional vector Brownian motion with covariance matrix

$$\Sigma := \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\eta} \\ \sigma_{\eta\varepsilon} & \Sigma_\eta \end{pmatrix}. \tag{2.8}$$

Let $z_t = (\Delta y_{t-1}, ..., \Delta y_{t-p}, \Delta x_{t+r}, ..., \Delta x_{t-q})'$

Assumption 2.2. $\sigma_u^2 > 0$ *and* $E(z_t z_t') > 0$

Remark 2.3. *Assumption 2.2 ensures that the series of interest,* $\{y_t\}$*, follows a unit root process under the null hypothesis of* $\delta = 0$. $E(z_t z_t)' > 0$ *ensures that the stationary regressors in* (2.4) *are asymptotically linearly independent, which is required along with Assumption 2.1(a) to ensure consistency of the least squares estimate of* δ.

2.3 Extant Recursive Test Procedures

Phillips et al. (2015) initially propose a univariate approach to testing for end-of-sample bubbles utilising the standard (non-covariate augmented) ADF regression given by;

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{k=1}^{p} \alpha_k \Delta y_{t-k} + e_t \tag{2.9}$$

performed on the full sample of data, where μ is the intercept and p is the number of lags of the dependent variable Δy_t. We denote the ADF test applied to the full sample as $ADF_0^1(p)$.

Full sample tests for explosive behaviour can be shown, however, to have very poor power to detect a short lived explosive episode in a series that otherwise follow a unit root process. As such, Phillips et al. (2015) consider test statistics that are functions of a sequence of ADF statistics applied to subsamples of the data. Specifically, if we denote an ADF test procedure performed on the subsample $t = \lfloor r_1 T \rfloor, ..., \lfloor r_2 T \rfloor$ as $ADF(p)_{r_1}^{r_2}$, then Phillips et al. (2015) propose the following test statistic to test for an explosive episode;

$$SADF := \sup_{r_2 \in [r_0, 1]} \{ADF(p)_0^{r_2}\}. \tag{2.10}$$

The SADF test is the supremum of right-tailed ADF statistics performed on a window of observations starting at $t = 1$ subject to a minimum sample size $\lfloor r_0 T \rfloor$. This recursive regression technique constitutes a powerful tool for detecting periodically collapsing explosive behaviour and can as well be utilised for confidence interval construction. One drawback of the SADF test is that it lacks power to detect end-of-sample explosive episodes that are arguably of most interest empirically, as more early observations relative to end-of-sample observations are used in its construction.

Motivated by this lack of power to detect end-of-sample explosive episodes Phillips et al. (2015) propose utilising the following test statistic to test for an end-of-sample explosive episode;

$$BSADF := \sup_{r_1 \in [0, 1 - r_0]} \{ADF(p)_{r_1}^{1}\}. \tag{2.11}$$

The BSADF test is the supremum of right-tailed ADF statistics computed on all subsamples ending at date $t = T$ subject to a minimum sample size $\lfloor r_0 T \rfloor$. The BSADF test is designed to detect end-of-sample explosive episodes and is as well shown by Phillips et al. (2015) to be useful for date stamping past explosive episodes. The BSADF test is particularly powerful when the bubble episode occurs at the end of the sample since the BSADF test is constructed in such a way that each subsample ADF test used in its construction will be computed using observations from the end-of-sample explosive regime.

Finally, Phillips et al. (2015) propose the GSADF test that its test statistic constructed from a sequence of ADF test statistics computed over all possible start and end dates, again, subject to a minimum sample size. The GSADF test of Phillips et al.

(2015) thus takes the form;

$$GSADF := \sup_{\substack{r_2 \in [r_0,1] \\ r_1 \in [0, r_2 - r_0]}} \{ADF(p)_{r_1}^{r_2}\}. \tag{2.12}$$

The GSADF test is a double-recursive unit root test, designed to test for the presence of one or more explosive episodes in a financial time series that are permitted to occur anywhere in the sample. As such, a rejection when using the GSADF test can only signal that a sample contains at least one explosive episode, but not where in the sample the explosive episode occurs.

2.4 Extant Covariate Augmented Unit Root Tests

Hansen(1995) proposes testing the null hypothesis of a unit root ($H_0 : \delta = 0$) in $\{y_t\}$ using the test t-statistic $t_{\hat{\delta}} = \frac{\hat{\delta}}{s(\hat{\delta})}$ where $s(\hat{\delta})$ is the standard error of $\hat{\delta}$ calculated from Equation (2.4). We refer to $t_{\hat{\delta}}$ as the $CADF(p, r, q)_0^1$ (*Covariate Augmented Dickey-Fuller*) test statistic where p, r, q denote the order of the lag polynomials $\alpha(L)$ and $b(L)$ specified in Equation (2.3).

The asymptotic distribution of the $t(\hat{\delta})$ test statistic under H_0 and appropriate critical values for a left-tailed test against the alternative of stationarity are provided in Hansen (1995), who shows that the inclusion of relevant covariates in the unit root test procedure leads to a substantial increase in power compared to a univariate approach. In the context of testing for explosive episodes we could simply utilise a right-tailed version of this test procedure and expect significant power gains relative to using a non covariate augmented full sample ADF test, however, it has been shown by *inter alia* Phillips et al. (2015) that the power of full sample based tests for explosivity lack power relative to newly proposed recursive test procedures. As such, we will now explore the possibility of utilising covariate augmented test statistics performed in a recursive manner.

2.5 Proposed Tests

A potential drawback of the tests of Phillips et al. (2015) outlined in Section 2.4 is that a univariate process is assumed. Given the complex relationships across multiple asset prices a multivariate approach could, potentially, be of much greater use. If a practitioner is interested in testing the null hypothesis of a unit root in $\{y_t\}$ against the alternative of explosivity for the entire sample then one could simply estimate regression (2.4) over the entire sample and perform a right tailed t-test of the null of $\delta = 0$ using critical values from Hansen (1995). Given the power improvements offered by the supremum ADF tests of Phillips et al. (2015) relative to the full sample ADF test in a univariate setting, however, we consider utilising the following test statistics instead;

$$CSADF := \sup_{r_2 \in [r_0, 1]} \{CADF(p, r, q)_0^{r_2}\}, \tag{2.13}$$

that is the supremum of covariate augmented Dickey-Fuller test statistics computed over all possible end dates for samples starting at time $t = 1$ (subject to a minimum sample size $r_0 T$),

$$CBSADF := \sup_{r_1 \in [0, 1-r_0]} \{CADF(p, r, q)_{r_1}^{1}\}, \tag{2.14}$$

which is the backward supremum of covariate augmented Dickey-Fuller test statistics computed over all possible start dates for samples ending at time $t = T$ (subject to a minimum sample size $\lfloor r_0 T \rfloor$) and

$$CGSADF := \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \{CADF(p, r, q)_{r_1}^{r_2}\}. \tag{2.15}$$

that is the generalised supremum of covariate augmented Dickey-Fuller test statistics computed over all possible start dates and end dates (subject to a minimum sample size $r_0 T$). Such an approach will be able to exploit both the inter dependencies between the series $\{y_t\}$ and any relevant covariates and the power gains associated with adopting a recursive estimation approach found by Phillips et al. (2015).

2.5.1 Limit Theory

In this section we outline the limiting null distribution of the $CGSADF$ test statistic, with the limiting null distributions of the $CSADF$ and CBSADF test statistics following as special cases of this result.

Theorem 2.1. *Let data be generated according to* (2.1) - (2.3) *and additionally let Assumptions 2.1 and 2.2 hold, then under the null hypothesis of no explosivity we have*

(a) If $d_t = 0$

$$CGSADF \xrightarrow{d} \sup_{\substack{r_2\in[r_0,1] \\ r_1\in[0,r_2-r_0]}} \frac{\int_{r_1}^{r_2} Q(s)dP(s)}{\left(\int_{r_1}^{r_2} Q(s)^2 ds\right)^{1/2}} \tag{2.16}$$

where $Q(s) = b(1)\Psi(1)B_\eta(s) + B_\varepsilon(s)$ *and* $P(s) = B_\varepsilon(s)/\sigma_\varepsilon$.

(b) If $d_t = \mu$

$$CGSADF \xrightarrow{d} \sup_{\substack{r_2\in[r_0,1] \\ r_1\in[0,r_2-r_0]}} \frac{\int_{r_1}^{r_2} Q^\mu(s)dP(s)}{\left(\int_{r_1}^{r_2} Q^\mu(s)^2 ds\right)^{1/2}} \tag{2.17}$$

where $Q^\mu(s) = Q(s) - \frac{1}{(r_2-r_1)}\int_{r_1}^{r_2} Q(t)dt$ *and* $P(s) = B_\varepsilon(s)/\sigma_\varepsilon$.

(c) If $d_t = \mu + \theta t$

$$CGSADF \xrightarrow{d} \sup_{\substack{r_2\in[r_0,1] \\ r_1\in[0,r_2-r_0]}} \frac{\int_{r_1}^{r_2} Q^\tau(s)dP(s)}{\left(\int_{r_1}^{r_2} Q^\tau(s)^2 ds\right)^{1/2}} \tag{2.18}$$

where $Q^\tau(s) = Q^\mu(s) - 12(r_2-r_1)^{-3}\left(s - \frac{(r_2-r_1)}{2}\right)\int_{r_1}^{r_2}\left(t - \frac{(r_2-r_1)}{2}\right)Q^\mu(t)dt$ *and* $P(s) = B_\varepsilon(s)/\sigma_\varepsilon$.

Remark 2.4. *The asymptotic distribution of the CSADF test follows directly from Theorem 5.1 by fixing* $r_1 = 0$*, whereas the asymptotic distribution of the CBSADF test obtains by fixing* $r_2 = 1$.

The proof of this result can be found in the Appendix.

2.5.2 Practical Implementation of tests

The asymptotic distribution of our proposed tests depends on the nuisance parameter ϱ which is the long-run correlation coefficient between the equation error and the covariate, with ϱ^2 measuring the relative contribution of the covariates Δx_t to the error term u_t. The coefficient ϱ^2 can take values between 0 (when the covariates fully explain the variability of the error term) and 1 (when the covariates have no explanatory power). We expect that the lower the long-run correlation between ε_t and u_t, that is for lower values of ϱ^2, the greater the power gains from the inclusion of the covariates into our proposed tests relative to univariate methods.

In practice the true value of ϱ^2 is unknown. We therefore propose selecting critical values for our proposed test procedures using a consistent non-parametric estimator of ϱ^2 given by

$$\hat{\varrho}^2 = \frac{\hat{\sigma}_{u\varepsilon}^2}{\hat{\sigma}_{\varepsilon}^2 \hat{\sigma}_u^2} \tag{2.19}$$

where

$$\hat{\Omega} = \begin{bmatrix} \hat{\sigma}_u^2 & \hat{\sigma}_{u\varepsilon} \\ \hat{\sigma}_{u\varepsilon} & \hat{\sigma}_{\varepsilon}^2 \end{bmatrix} = \sum_{k=-M}^{M} w(k/M) T^{-1} \sum_t \hat{\gamma}_{t-k} \hat{\gamma}_t' \tag{2.20}$$

and $\hat{\gamma}_t = (\hat{u}_t, \hat{\varepsilon}_t)'$ are least squares estimators of $\gamma_t = (u_t, \varepsilon_t)'$ from the full sample estimation of Equations (2.1) - (2.3) The function w is the Parzen kernel function that produces positive semidefinite covariance matrices;

$$w(x) = \begin{cases} 1 - 6x^2 + 6|x|^3 & \text{for } 0 \leq |x| \leq 1/2, \\ 2(1 - |x|)^3 & \text{for } 1/2 \leq |x| \leq 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2.21}$$

and M is an automatic bandwidth estimator of Andrews (1991) that grows slowly with the sample size;

$$M = 2.6614(\alpha(2)T)^{1/5}, \tag{2.22}$$

where $\alpha(2)$ is a function of an unknown spectral density matrix as discussed in Andrews (1991). The number of lags of the dependent variable Δy_t as well as the number of lags, q, and leads, r, of the covariate Δx_t are chosen using the Bayesian Information Criterion (BIC).

The methodology described above can be shown to perform well when used as part of a full sample testing process, but we will show that this methodology has poor size control in a recursive sub-sample framework, mainly due to imprecise estimation of the nuisance parameter ϱ^2. We therefore consider bootstrap implementation of our proposed test procedures in order to better control size in a recursive estimation framework.

2.5.3 Bootstrap Unit Root Tests with Covariates

To better control size in finite samples when performing tests in a recursive framework we propose bootstrap implementation of our proposed covariate augmented test procedures. Following Chang, Sickles and Song (2017) we utilise the following bootstrap algorithm.

Algorithm 1. ***(Bootstrap CADF tests)***

Step 1: *Compute $u_t = \Delta y_t$ and estimate the following regression by OLS;*

$$u_t = \sum_{k=1}^{p} \tilde{a}_k u_{t-k} + \sum_{k=-r}^{q} \tilde{\beta}_k' \Delta x_{t-k} + \tilde{\varepsilon}_t. \tag{2.23}$$

with p, q and r chosen using the Bayesian Information Criterion (BIC).

Step 2: *Estimate the following regression using the Yule-Walker method*[10]*;*

$$\Delta x_{t+r+1} = \tilde{\Phi}_{1,n} \Delta x_{t+r} + ... + \tilde{\Phi}_{l,n} \Delta x_{t+r-l+1} + \tilde{\eta}_t. \tag{2.24}$$

Step 3: *Define $\tilde{\xi}_t = (\tilde{\varepsilon}_t, \tilde{\eta}_t)$ where $\tilde{\varepsilon}_t$ and $\tilde{\eta}_t$ are the fitted residuals obtained from Equations (2.23) and (2.24) and generate bootstrap samples (ξ_t^*) by resampling from the centred distribution of $\tilde{\xi}_t$;*

$$\left(\tilde{\xi}_t - \frac{1}{n}\sum_{t=1}^{n} \tilde{\xi}_t\right)_{t=1}^{n}.$$

Step 4: *Construct recursively the bootstrap samples of the stationary covariate (Δx_t^*) according to;*

$$\Delta x_{t+r+1}^* = \tilde{\Phi}_{1,n} \Delta x_{t+r}^* + ... + \tilde{\Phi}_{l,n} \Delta x_{t+r+1-\ell}^* + \tilde{\eta}_t^* \tag{2.25}$$

where the ℓ-initial values of Δx_t^* are set equal to zero.

Step 5: *Construct the bootstrap sample $\{v_t^*\}$ according to;*

$$v_t^* = \sum_{k=-r}^{q} \tilde{\beta_k}' \Delta x_{t-k}^* + \varepsilon_t^* \tag{2.26}$$

using the OLS estimates $\tilde{\beta}_k$, $-r \leq k \leq q$ from Equation (2.23).

Step 6: *Generate recursively the bootstrap samples of the error term $\{u_t^*\}$ from;*

$$u_t^* = \tilde{\alpha}_1 u_{t-1}^* + ... + \tilde{\alpha}_p u_{t-p}^* + v_t^* \tag{2.27}$$

where we set the initial values $u_0^*, ..., u_{-(p-1)}^* = 0$ and $\tilde{\alpha}_k$, $1 \leq k \leq p$ are estimated from Equation (2.23).

Step 7: *Generate $\{y_t^*\}$ from cumulating the bootstrap values of $\{u_t^*\}$;*

[10]Brockwell and Davis (1991) suggest to utilise the Yale-Walker method for estimating ARMA(p,q) models to obtain good estimates in small samples and (or) when $q = 0$ to ensure stationarity.

$$y_t^* = y_{t-1}^* + u_t^* = y_0^* + \sum_{k=1}^{t} u_k^* \tag{2.28}$$

where we set $y_0^* = 0$.

Step 8: *A covariate Augmented Dickey-Fuller unit root test statistic calculated using the observations* $\{y_t^*\}$ *from* $t = \lfloor r_1 T \rfloor, ..., \lfloor r_2 T \rfloor$, $(CADF^*)_{r_1}^{r_2}$, *can then be calculated from the following regression;*

$$\Delta y_t^* = \alpha y_{t-1}^* + \sum_{k=-r}^{p} \alpha_k \Delta y_{t-k}^* + \sum_{k=-r}^{q} \beta_k' \Delta x_{t-k}^* + \varepsilon_t^* \tag{2.29}$$

Remark 2.5. *The algorithm above is outlined for the case where no deterministic components are allowed for in the data. If a constant is to be allowed for in the data then the series* y_t *and* Δx_t *should be replaced by their demeaned counterparts. Likewise, if a constant and trend are to be allowed for in the data then* y_t *and* Δx_t *should first be demeaned and detrended.*

We propose utilising the CADF*(p,r,q) test statistic in place of the standard CADF(p,r,q) to deliver a test with controlled finite sample size when estimation is performed in a recursive framework. Our proposed tests are, therefore, given by;

$$CSADF^* := \sup_{r \in [r_0, 1]} \{CADF^*(p,r,q)_0^r\} \tag{2.30}$$

$$CBSADF^* := \sup_{r \in [0, 1-r_0]} \{CADF^*(p,r,q)_r^1\} \tag{2.31}$$

$$CGSADF^* := \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \{CADF^*(p,r,q)_{r_1}^{r_2}\} \tag{2.32}$$

The benefit of the bootstrap procedure is that the impact of ϱ^2 on the critical values of the proposed tests is now modelled explicitly. The bootstrap CADF test deals with the nuisance parameter dependency problem and will be shown to better control size in finite samples compared to the non-bootstrap version of the tests. In what follows we will focus on the performance of the CBSADF* test as this test is constructed to detect end-of-sample and ongoing explosive episodes which are arguably of most interest to practitioners.

2.5.4 Bootstrap Limiting Distribution

In this section we outline the bootstrap limiting null distribution of the $CGSADF^*$ test statistic, with the limiting null distributions of the $CSADF^*$ and CBSADF* test statistics following as special cases of this result.

Theorem 2.2. *Let data be generated according to* (2.1) -(2.3) *and additionally let Assumptions 2.1 and 2.2 hold, then under the null hypothesis of no explosivity we have*

(a) If $d_t^* = 0$

$$CGSADF^* \xrightarrow{d^*} \sup_{\substack{r_2 \in [r_0,1] \\ r_1 \in [0, r_2 - r_0]}} \frac{\int_{r_1}^{r_2} Q(s) dP(s)}{\left(\int_{r_1}^{r_2} Q(s)^2 ds \right)^{1/2}} \tag{2.33}$$

where $Q^(s) = b(1)\Psi(1)B_\eta(s) + B_\varepsilon(s)$ *and* $P(s) = B_\varepsilon(s)/\sigma_\varepsilon$.

(b) If $d_t^* = \mu^*$

$$CGSADF^* \xrightarrow{d^*} \sup_{\substack{r_2 \in [r_0,1] \\ r_1 \in [0, r_2 - r_0]}} \frac{\int_{r_1}^{r_2} Q^\mu(s) dP(s)}{\left(\int_{r_1}^{r_2} Q^\mu(s)^2 ds \right)^{1/2}} \tag{2.34}$$

where $Q^\mu(s) = Q(s) - \frac{1}{(r_2 - r_1)} \int_{r_1}^{r_2} Q(t) dt$ *and* $P(s) = B_\varepsilon(s)/\sigma_\varepsilon$.

(c) If $d_t^* = \mu^* + \theta^* t$

$$CGSADF^* \xrightarrow{d^*} \sup_{\substack{r_2 \in [r_0,1] \\ r_1 \in [0, r_2 - r_0]}} \frac{\int_{r_1}^{r_2} Q^\tau(s) dP(s)}{\left(\int_{r_1}^{r_2} Q^\tau(s)^2 ds \right)^{1/2}} \tag{2.35}$$

where $Q^\tau(s) = Q^\mu(s) - 12(r_2 - r_1)^{-3} \left(s - \frac{(r_2 - r_1)}{2} \right) \int_{r_1}^{r_2} \left(t - \frac{(r_2 - r_1)}{2} \right) Q^\mu(t) dt$ *and* $P(s) = B_\varepsilon(s)/\sigma_\varepsilon$.

Remark 2.6. *The asymptotic distribution of the* $CSADF^*$ *test follows directly from Theorem 2.2 by fixing* $r_1 = 0$, *whereas the asymptotic distribution of the CBSADF* test obtains by fixing* $r_2 = 1$.

The proof of this result can be found in the Appendix.

2.6 Finite Sample Simulations

In this section we examine the finite sample size and power properties of our proposed tests relative to the extant tests of Phillips et al. (2015). In order to do so, data were simulated according to the following data generating process that allows for a single covariate;

$$y_t = \phi_t y_{t-1} + u_t,\ t = 1, ..., T \tag{2.36}$$

with

$$u_t = \alpha_1 u_{t-1} + v_t \tag{2.37}$$

$$v_t = \beta \Delta x_t + \varepsilon_t, \tag{2.38}$$

$$\Delta x_{t+1} = \lambda \Delta x_t + \eta_t, \tag{2.39}$$

and where

$$\xi_t = \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{\varepsilon\eta} \\ \sigma_{\eta\varepsilon} & 1 \end{bmatrix} \right). \tag{2.40}$$

The perfomance of all tests will depend on the correlation between the error term v_t and the stationary covariate Δx_t and therefore on λ and β as these two coefficients determine the degree of correlation between the series of interest y_t and the covariate Δx_t.

Following Chang, Sickles and Song (2017) we examine values of λ and β between -0.8 and 0.8 and set $\alpha_1 = 0.2$ and $\sigma_{e\eta} = 0.4$. As in Hansen (1995) we simulate series of length $T + 100$ and drop the first 100 observations to eliminate any start-up effects. The minimum window size, r_0 for all recursive test procedures is chosen as;

$$r_0 = (0.01 + \frac{1.8}{\sqrt{T}}) * T. \tag{2.41}$$

The choice of lag lengths is of particular importance and critically affects the finite sample performance of the tests. We use the Bayesian Information Criterion (BIC) to decide on the optimal lag length of the dependent variable Δy_t as well as the optimal lag and lead lengths of the stationary covariate Δx_t, with the maximum lag set to four for the dependent variable Δy_t and the maximum lag and lead for the covariate Δx_t set to two.

All simulations that follow were conducted in GAUSS 17 using $2,000$ Monte Carlo replications and 999 bootstrap replications. All tests are performed at a nominal 5% level of significance. In the bootstrap algorithm, a constant is allowed for in the data and therefore the series y_t and Δx_t are replaced by their demeaned counterparts.

2.6.1 Empirical Size

To assess the size properties of our proposed tests, data were generated according to Equations (2.36)-(2.40) with $\phi_t = 1$ for the full sample $t = 1, ..., T$ after discarding the first 100 observations to eliminate any start-up effects. We report the empirical size of the BSADF test of Phillips et al. (2015), the CBSADF$\hat{\varrho}^2$ test (i.e. the CBSADF test based on the estimated long-run correlation coefficient squared), the CBSADFϱ^2 test (i.e. the CBSADF test based on the true value of the long-run correlation coefficient squared), the bootstrap version of the BSADF test of Phillips et al. (2015) and our proposed CBSADF* test, for a range of values of $\beta, \lambda \in [-0.8, 0.8]$ along with the true value of ϱ^2 for each parameter setting. Table 1 reports the finite sample size of all tests for a sample size of $T \in [100, 250, 400]$.

In Table 1 the non-bootstrap BSADF test and both the non-bootstrap CBSADF tests exhibit severe size distortions across most of the scenarios considered, with the size of these tests exceeding or being below the nominal 5% level of significance in a number of scenarios. On the other hand, the size distortions exhibited by the CBSADF$\hat{\varrho}^2$ test in some of the scenarios can be attributed to the imprecise estimation of the nuisance parameter ϱ^2. The limit null distribution and critical values of the CBSADF$\hat{\varrho}^2$ are a function of $\hat{\varrho}^2$ and, therefore, imprecise estimation of ϱ^2 leads to incorrect critical values being utilised, resulting in inevitable size distortions, see for instance when $\beta = -0.8$ and $\lambda = 0.8$.

To continue with, the bootstrap based BSADF* test displays much better size control for most of the designs compared to the non-bootstrap BSADF test of Phillips et al. (2015). When $T = 100$ the bootstrap based CBSADF* test displays some modest oversize across some of the scenarios considered, exhibiting size ranging from 0.05 to 0.09. We attribute these size distortions to the small size of the sample.

Increasing the sample size to $T \in [250, 400]$ in Table 1, improves the size performance of our proposed bootstrap based CBSADF* test proving that the size distortions are a small sample issue. The size performance of the bootstrap CBSADF* test is significantly improved for larger samples as can been seen in Table 1, when $T = 250$ while when $T = 400$, our proposed test displays even better finite sample size properties indicating that the oversize exhibited by our proposed bootstrap CBSADF* test eliminated for larger samples. The CBSADF* test displays excellent size control across all scenarios, with some size distortions exhibited by this test in the smaller sample size of $T = 100$ greatly improved in the larger sample size of $T = 250$ and $T = 400$.

Overall, it can be seen that both the bootstrap BSADF* and CBSADF* tests control size to a similar or greater degree than the respective non-bootstrap based BSADF and CBSADF tests, with the CBSADF* test displaying some modest size distortions for small samples, that they would be of little use empirically in the scenarios considered. Whilst the bootstrap BSADF* displays greater size control than its non-bootstrap counterpart, the CBSADF* test, has the best overall size control, with oversize exhibited by this test in smaller sample sizes, entirely eliminated when a larger sample size is considered.

2.6.2 Empirical Power

We now proceed to examine the power of our proposed tests relative to extant tests. To do so, data were generated according to Equations (2.36)-(2.40) with $T = 250$ under the alternative hypothesis of an end-of-sample explosive episode by setting $\phi_t = 1$ for $t = 1, ..., 200$, and $\phi_t = \phi > 1$ for $t = 201, ..., 250$. The series $\{y_t\}$ follows a unit root process for the first 200 observations and is then subject to (potential) explosive behaviour over the remaining 50 observations. Due to the severe oversize exhibited by the BSADF test and the CBSADF test in some scenarios, we report power results for the BSADF* and CBSADF* tests as well as the size-adjusted BSADF and CBSADF tests.

The finite sample power of the BSADF* and CBSADF* tests and the size-adjusted BSADF and CBSADF tests is computed over a grid of 20 values of ϕ from $\phi = 1.00$ to $\phi = 1.05$ for each of the eight pairings of β and λ discussed previously. Figures 1 and 2 report finite sample power curves for the BSADF* and CBSADF* tests across each of the scenarios considered (pairings of β and λ).

In all scenarios the power of all tests is increasing monotonically in ϕ. In most of the scenarios the CBSADF* test displays greater power overall compared to the BSADF* test or the size-adjusted BSADF and CBSADF tests, although for some designs the power of the CBSADF* test is lower than that of the BSADF* test for lower values of ϕ mainly due to the oversize exhibited by the BSADF* test in most of the scenarios, see for example Figure 2 (c). In general, however, the power of the CBSADF* test quickly exceeds that of the BSADF* test and the size-adjusted BSADF and CBSADF tests as

ϕ increases.

The CBSADF* test displays much greater power than the BSADF* test, with the power differential between the two tests reaching almost 20% and an even greater power differential between the CBSADF* and the size-adjusted BSADF test, for a given value of ϕ as can be seen in Figure 1 (a). Finally, there is no significant contribution of the covariate in Figure 2 (d) since all four tests (i.e. size-adjusted BSADF test, size-adjusted CBSADF test, bootstrap BSADF* test and CBSADF* test) appear to have similar power properties.

Overall, we argue that both the bootstrap BSADF* and CBSADF* tests show better size control than their respective non-bootstrap (BSADF and CBSADF) tests across most of the parameter values while offering significant power gains, with the best performance given by the CBSADF* test. Arguably, the inclusion of covariates in the CBSADF* test leads to greater power performance relative to the BSADF* test in finite samples, as well as offering slightly improved size control. We, therefore, recommend utilising the CBSADF* test in practice as it offers the best overall size control and power properties amongst the tests considered.

2.7 Empirical Application

Following Phillips et al. (2015) empirical application, we consider the real S&P 500 stock price index and the real S&P 500 stock price index dividend over the period January 1959 to June 2018 at a monthly frequency, constituting 714 observations.[11] We utilise the same dataset with Phillips et al. (2015) as it contains multiple historical bubble episodes and we estimate the present value of the real price-dividend ratio which is the real S&P 500 stock price index over the real S&P 500 stock price index dividend as outlined in Phillips et al. (2015).

According to Shiller (2015) bonds are related to asset bubbles as when long-term interest rates decrease, bond prices increase creating enthusiasm in a similar way as in the stock market. Vogel (2010) argues that during the end of the double recessions of the 80s, the bond market had reached historical low levels as a consequence of the tight monetary policy and high interest rates FED implemented to deal with double-digit inflation. The "bond market conundrum", the inability of the monetary policy to

[11]Both obtained from Robert Shiller's website: http://www.econ.yale.edu/ shiller/data.htm

affect long-term bond yields has been widely acknowledged in the literature (see *inter alia* Evanoff, Kaufman and Malliaris 2012, Bernanke 2005). Furthermore, monetary tightening might not be able to affect long-term interest rates if the FED does not increase the federal funds rate enough prior to the bubble episode, as a great increase in long term interest rates can mitigate the episode (Taylor 2007). As a long-term interest rate, we choose the Ten-Year Treasury Constant Maturity Yield which is published by the Federal Reserve Board and it is calculated based on the daily yield curve for non-inflation-indexed Treasury securities, all adjusted to the equivalent of a 10-year maturity and it is based on the closing market bid yields on actively traded Treasury securities in the over-the-counter market (Federal Reserve statistical release 2016).

The CBOE S&P 100 volatility index (initially VIX, renamed to VXO in 2003) measures market's expectation of 30-day volatility and it is constructed by using the implied volatilities on the S&P 100 index options (OEX). OEX options were the standard index options traded in the domestic stock market, however the trading activity on EOX options started decreasing as more and more investors started trading using S&P 500 index options instead, contributing to the introduction of the CBOE volatility index (VIX) in 2003 when the CBOE S&P 100 volatility index had its symbol changed from VIX to VXO. In their seminal work, Fleming, Ostdiek and Whaley (1995) provide evidence in support of the argument that there is a tendency of the VIX (which it was renamed to VXO in 2003) to rise after large sell-offs and fall after large rallies. In 1998, during the dot-com bubble episode, VIX appeared having a quite wide range as 90% of the VIX levels were between 18.57% and 42.74% whereas after the collapse of the dot-com bubble the range according to Whaley (2000) narrowed to 11% (20% - 31%). The VIX index is known as *fear index* as it reflects investors expectations about future volatility as well as their willingness to pay in terms of implied volatility to hedge their stock portfolios (Whaley 2000).

For all the above, we utilise as covariates the Moody's Seasoned Aaa Corporate Bond Yield[12] as well as the Moody's Seasoned Baa Corporate Bond Yield[13] that both cover the period between January 1959 and June 2018 (714 observations), the Chicago Board Options Exchange Volatility Index (CBOE VXO)[14] from January 1986 to June 2018

[12]Retrieved from: https://fred.stlouisfed.org/series/AAA
[13]Retrieved from: https://fred.stlouisfed.org/series/BAA
[14]Retrieved from: https://finance.yahoo.com/quote/%5EVXO/

(390 observations) and the Ten-Year Treasury Constant Maturity Yield (GS10)[15] over the period March 1970 to June 2018 (580 observations). All covariates are sampled at a monthly frequency.

Before we include the covariates into our proposed CBSADF* test it is important that we firstly ensure the stationarity of the covariates. In Table 2 we report the full sample ADF test statistics for all four covariates together with the finite sample left-tailed critical values for their corresponding sample size. We choose to include a constant in the ADF regression for the covariates and the lags of Δx_t is set to zero as suggested by the Bayesian Information Criterion (BIC).

Performing a unit root test on the Aaa and Baa Corporate Bond Yields, we conclude that we cannot reject the null hypothesis of a unit root at 5% level of significance and therefore we are including first differences of the series in the CADF regression. Testing the Ten-Year Treasury Constant Maturity Yield for a unit root we conclude that we cannot reject the null hypothesis and therefore we utilise this covariate in differences as well. Finally, although the Volatility Index (VXO) appears to be stationary, we strongly reject the null hypothesis of a unit root at 5% level of significance, it is commonly argued in the literature that asset price volatility is shown to be fractionally integrated as it retains long memory (see *inter alia* Bollerslev and Mikkelsen 1996, Parke 1999 and Fantazzini 2011). For this reason, the Volatility Index (VXO) is used in first differences.

In the bootstrap algorithm, a constant is allowed for in the data and therefore the series y_t and Δx_t are replaced by their demeaned counterparts. To continue, we apply the BSADF* and CBSADF* tests as outlined earlier in sections 2.4 and 2.5. In particular, we consider the price-dividend ratio as the dependent variable regressed on its lags and leads and lags of the covariates, namely the Aaa corporate bond yield, Baa corporate bond yield, ten-year treasury rate and volatility index (VXO). Furthermore, we use the Bayesian Information Criterion (BIC) to determine the lags of the differenced dependent variable. If a practitioner is interested in testing for bubbles in real time then one would only consider lags and not leads of the potential covariates. Since we are referring to past bubble episodes we choose to make use of leads as well, but our proposed tests can be equally useful on a real-time basis when using leads might not be feasible.

We utilise all three covariates to identify bubble episodes in historical prices applying

[15] Retrieved from https://fred.stlouisfed.org/series/GS10

our proposed CBSADF* test together with the bootstrap version of the BSADF test of Phillips et al. (2015), i.e BSADF* test. Particularly, we use the Aaa corporate bond yield and Baa corporate bond yield for the period January 1959 to June 2018, the ten-year treasury yield from March 1970 to June 2018 and the volatility index (VXO) from January 1986 to June 2018. Results are summarised in Table 3.

In addition, we compute right-tailed finite sample critical values for both tests using 999 bootstrap replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regression.

To investigate the accuracy of our proposed tests to detect bubble episodes in empirical data series, we follow Phillips et al. (2015) who perform a (pseudo) real-time bubble monitoring exercise on the present value of the real S&P 500 price-dividend ratio and apply a date-stamping strategy to test for the presence of explosive behaviour. Particularly, we estimate both the BSADF* and CBSADF* test statistics in a recursive framework and we identify the origination date of the bubble episode as the first chronological observation of which the test statistic is larger than the simulated finite sample critical value therefore rejecting the null hypothesis of a unit root. On the same spectrum, the termination date of the bubble episode is defined as the first chronological observation of which test statistic becomes less than the simulated finite sample critical value.

Table 3 presents three periods of bubble episodes (origination and termination dates) as identified by the BSADF* test and our proposed CBSADF* test across three covariates. We focus on the performance of the BSADF* and CBSADF* tests as detecting explosive episodes that occur at the end of the sample in real-time, can be useful to regulators, policy makers and central banks. Both the BSADF* and CBSADF* tests identify two periods of exuberance; namely the Black Monday of October 1987 and the dot-com bubble of 2000-2001. We present those explosive episodes that are detected earlier by the proposed CBSADF* test compared to the BSADF* test.

We report the origination and termination dates of two bubble episodes as identified by the BSADF* test and our proposed CBSADF* test across the four different covariates (Table 3). In particular, the BSADF* test identifies February 1987 as the origination date of the Black Monday episode, five months prior to the crash, whereas when the Aaa and Baa corporate bond yields are used as covariates independently, our proposed CBSADF* test detects the bubble episode eleven months earlier (March 1983) compared to the BSADF* test (February 1987). Furthermore, as it can be seen in Table 3, when utilising both the Aaa and Baa corporate bond yields as covariates independently, the CBSADF* test identifies the origination date of the dot-com bubble episode in July 1995, two months earlier than the BSADF* test (September 1995).

In a similar way, when the ten-year treasury rate is utilised as a covariate the proposed CBSADF* test identifies the origination date of the Black Monday episode in June 1986, nine months earlier compared to the BSADF* test (February 1987) and the origination date of the dot-com bubble episode in July 1995, while the BSADF* test identifies the origination date of the same episode two months later (September 1995). Finally, the BSADF* test proposes that the dot-com bubble episode originated in September 1995, whereas when the volatility index (VXO) is utilised as a covariate, the CBSADF* identifies the origination date two months earlier, in July 1995.

To conclude, our proposed CBSADF* test has identified the Black Monday of October 1987 as well as the dot-com bubble episode earlier compared to the BSADF* test, reflecting the superior power of the CBSADF* test to quickly detect bubble episodes. The CBSADF* test seems to be able to detect the bubble episodes earlier compared to the BSADF* test for all covariates. Finally, our empirical results are in accordance to the theoretical evidence presented in section 2.6, where we argue that the CBSADF* test shows better size control and power performance over the BSADF* test and we provide empirical evidence that the inclusion of covariates in the CBSADF* test leads to greater power properties relative to the BSADF* test in finite samples and therefore, recommend utilising the CBSADF* test in practice.

2.8 Conclusion

This chapter provides theoretical and empirical evidence that the inclusion of relevant covariates in the conventional Augmented Dickey Fuller test regression leads to improved size control, while offering significant power gains when an end-of-sample explosive episode is present. Specifically, we investigate whether the size and power properties of the tests of Phillips et al. (2015) can be enhanced by the inclusion of information in related time series that traditional univariate unit root tests tend to ignore.

When working only in a univariate framework, the choice to examine a variable in isolation can be rather costly in terms of power since ignoring any correlation with other relevant series leads to unnecessarily high standard errors when constructing the Dickey-Fuller test statistic, leading to significant power losses. Our proposed test is applied using a recursive window sequence as suggested by Phillips et al. (2015) to detect explosive episodes that occur at the end of the sample.

To deal with potential bias we apply a bootstrap version of the proposed test ensuring the asymptotic validity of the critical values drawn from the bootstrap distribution of the test. We concentrate on the case where the explosive episode takes place at the end of the sample as the detection of ongoing bubbles is of most importance to practitioners, with the tests being useful in terms of date stamping past bubbles as well.

Simulation results show that the proposed tests have improved size and power properties in finite samples relative to extant tests. In particular, the CBSADF* test suffers less severe size distortions compared to conventional tests that do not utilise a bootstrap procedure or omit relevant covariates, whilst displaying significantly better power properties.

Empirical work explores the effectiveness of the proposed tests as early warning mechanisms, informing policy makers of a bubble episode that might occur, which should provide evidence that the tests might be useful to structuring macroprudential policy. Specifically, we examine whether our proposed tests would have detected known past bubbles in the S&P 500 price dividend series before the tests of Phillips et al. (2015) when used as an early warning mechanism, utilising the Moody's Seasoned Aaa and Baa Corporate Bond Yields, the Ten-Year Treasury Rate and the Volatility Index as covariates. The superiority of our proposed test is reflected on the earlier detection of two major explosive episodes: Black Monday of October 1987 and the dot-com bubble.

2.9 Tables

Table 1: Finite Sample Size

$T = 100$							
β	λ	$BSADF$	$CBSADF_{\hat{\varrho}^2}$	$CBSADF_{\varrho^2}$	$BSADF^*$	$CBSADF^*$	ϱ^2
0.8	0.8	0.075	0.065	0.074	0.076	0.090	0.335
0.5	0.8	0.081	0.058	0.064	0.111	0.079	0.432
-0.5	0.8	0.058	0.233	0.108	0.113	0.073	0.000
-0.8	0.8	0.075	0.296	0.227	0.097	0.075	0.026
0.8	0.5	0.029	0.043	0.048	0.055	0.061	0.556
0.5	0.5	0.029	0.040	0.041	0.064	0.055	0.700
-0.5	0.5	0.012	0.083	0.056	0.077	0.050	0.300
-0.8	0.5	0.015	0.148	0.083	0.079	0.057	0.057
$T = 250$							
β	λ	$BSADF$	$CBSADF_{\hat{\varrho}^2}$	$CBSADF_{\varrho^2}$	$BSADF^*$	$CBSADF^*$	ϱ^2
0.8	0.8	0.038	0.057	0.059	0.065	0.062	0.335
0.5	0.8	0.051	0.049	0.051	0.092	0.067	0.432
-0.5	0.8	0.042	0.243	0.109	0.111	0.050	0.000
-0.8	0.8	0.050	0.335	0.247	0.103	0.046	0.026
0.8	0.5	0.009	0.037	0.039	0.040	0.051	0.556
0.5	0.5	0.012	0.029	0.028	0.057	0.056	0.700
-0.5	0.5	0.005	0.086	0.054	0.061	0.044	0.300
-0.8	0.5	0.008	0.163	0.079	0.058	0.043	0.057
$T = 400$							
β	λ	$BSADF$	$CBSADF_{\hat{\varrho}^2}$	$CBSADF_{\varrho^2}$	$BSADF^*$	$CBSADF^*$	ϱ^2
0.8	0.8	0.033	0.052	0.056	0.081	0.047	0.335
0.5	0.8	0.049	0.045	0.048	0.103	0.054	0.432
-0.5	0.8	0.041	0.262	0.124	0.128	0.050	0.000
-0.8	0.8	0.048	0.354	0.276	0.109	0.052	0.026
0.8	0.5	0.009	0.032	0.033	0.051	0.043	0.556
0.5	0.5	0.012	0.026	0.026	0.060	0.045	0.700
-0.5	0.5	0.004	0.076	0.046	0.055	0.033	0.300
-0.8	0.5	0.008	0.170	0.084	0.063	0.048	0.057

Data generated according to $y_t = y_{t-1} + u_t$ with $u_t = \alpha_1 u_{t-1} + v_t$, $v_t = \beta \Delta x_t + \varepsilon_t$ and $\Delta x_{t+1} = \lambda \Delta x_t + \eta_t$, where $\xi_t = \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{\varepsilon\eta} \\ \sigma_{\eta\varepsilon} & 1 \end{bmatrix} \right)$ $\alpha_1 = 0.2$ and $\sigma_{\eta\varepsilon} = \sigma_{\varepsilon\eta} = 0.4$.

Note: All simulations were conducted in GAUSS 17 using 2,000 Monte Carlo replications and 999 bootstrap replications. All tests are performed at a nominal 5% level of significance.

Table 2: Full Sample Unit Root Test

Panel A:	**Sample size**	**ADF**
	T	test statistic
Aaa Corporate Bond Yield	714	-1.090
Baa Corporate Bond Yield	714	-0.981
10-Year Treasury Rate	580	-0.916
Volatility Index (VXO)	390	-5.890

Panel B:	**finite sample critical values**		
Sample size	1%	5%	10%
$T = 714$	-3.400	-2.851	-2.567
$T = 580$	-3.400	-2.833	-2.539
$T = 390$	-3.430	-2.874	-2.558

Note: Finite sample left-tailed critical values of ADF test are obtained from Monte Carlo simulations with $10,000$ replications. A constant has been included in the ADF regression and the lags of Δx_t is set to zero as suggested by the Bayesian Information Criterion (BIC).

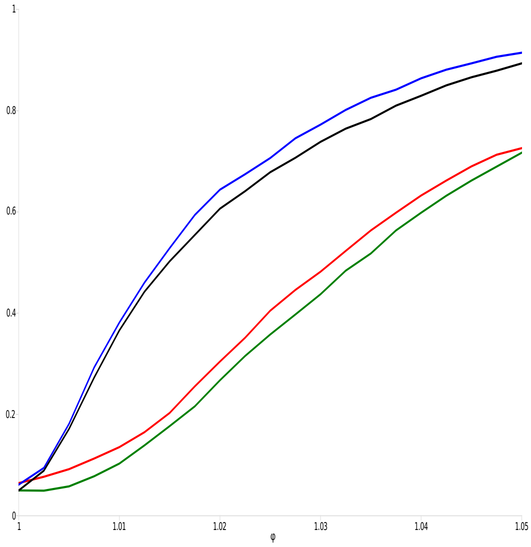
Table 3: Bubble Date Stamping

Covariate	**BSADF***	**CBSADF***
Aaa Corporate Bond Yield	1987M2 - 1987M9	1986M3 - 1987M12
	1995M9 - 2001M7	1995M7 - 2001M5
Baa Corporate Bond Yield	1987M2 - 1987M9	1986M6 - 1987M12
	1995M9 - 2001M7	1995M7 - 2001M5
10-Year Treasury Rate	1987M2 - 1987M9	1986M6 - 19871M2
	1995M9 - 2001M7	1995M7 - 2001M5
Volatility Index (VXO)	1995M9 - 2001M7	1995M7 - 2001M2

Note: Bubble date stamping application on the real S&P 500 stock price index and the real S&P 500 stock price index dividend over the period January 1959 to June 2018 at a monthly frequency, constituting 714 observations. We utilise the same dataset with Phillips et al. (2015) as it contains multiple historical bubble episodes and we estimate the present value of the real price-dividend ratio which is the real S&P 500 stock price index over the real S&P 500 stock price index dividend as outlined in Phillips et al. (2015). As covariates we utilise the Moody's Seasoned Aaa Corporate Bond Yield as well as the Moody's Seasoned Baa Corporate Bond Yield that both cover the period between January 1959 and June 2018 (714 observations), the Ten-Year Treasury Constant Maturity Yield (GS10) over the period March 1970 to June 2018 (580 observations) and the Chicago Board Options Exchange Volatility Index (CBOE VXO) from January 1986 to June 2018 (390 observations). All covariates are sampled at a monthly frequency and the present value of the real price-dividend ratio is equal to 100 at the beginning of the sample as in Phillips et al. (2015).

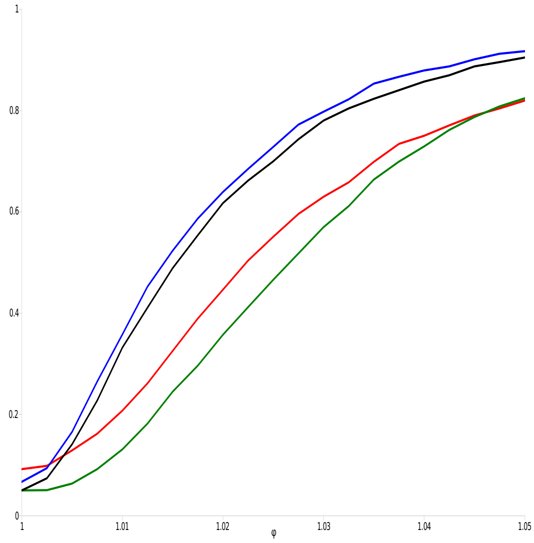
2.10 Figures

Figure 1: Finite Sample Power



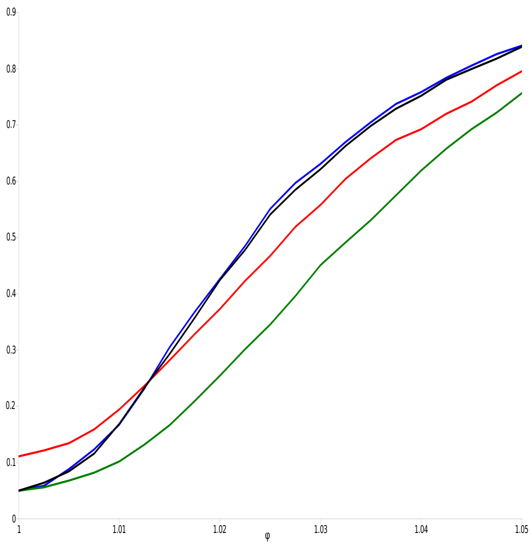
(a)

$\beta = 0.8 \quad \lambda = 0.8$



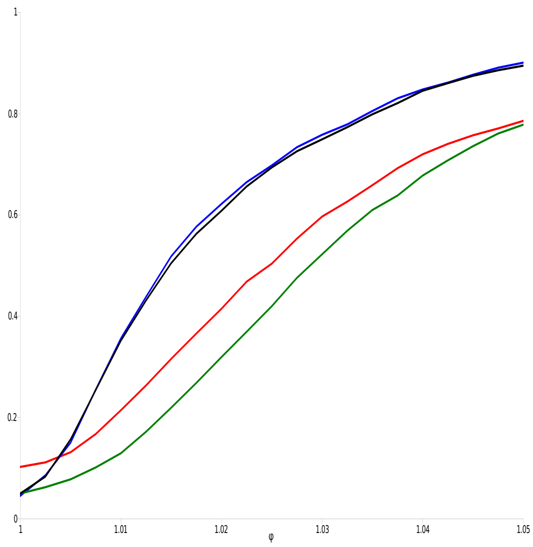
(b)

$\beta = 0.5 \quad \lambda = 0.8$



(c)

$\beta = -0.5 \quad \lambda = 0.8$



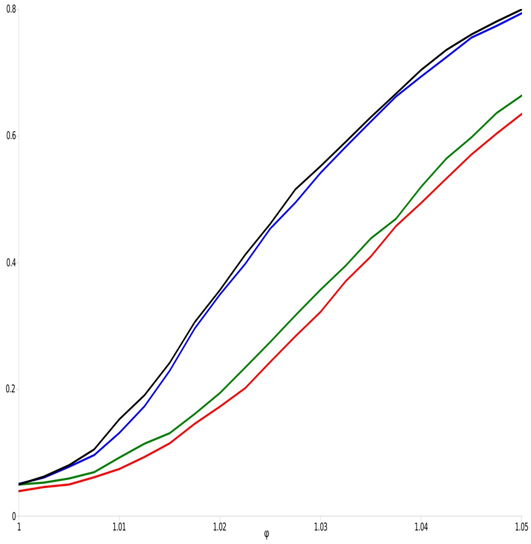
(d)

$\beta = -0.8 \quad \lambda = 0.8$

BSADF* test: —— CBSADF* test: ——
size-adjusted BSADF test: —— size-adjusted CBSADF test: ——

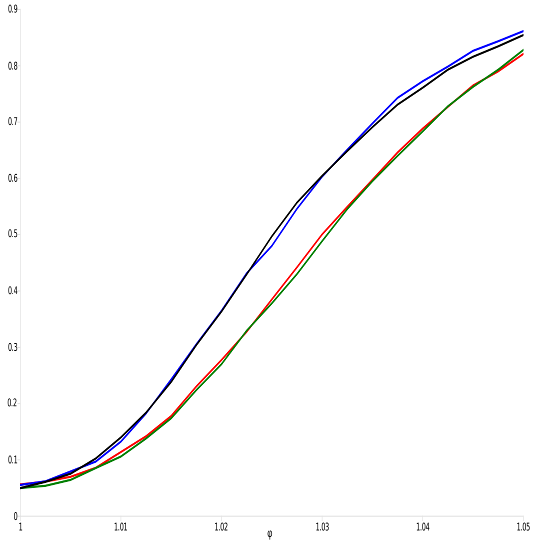
Data generated according to $y_t = y_{t-1} + u_t,\ t = 1, ..., 200$ and $y_t = \phi_t y_{t-1} + u_t, t = 201, ..., 250$ with $u_t = \alpha_1 u_{t-1} + v_t$, $v_t = \beta \Delta x_t + \varepsilon_t$ and $\Delta x_{t+1} = \lambda \Delta x_t + \eta_t$, where $\xi_t = \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{\varepsilon\eta} \\ \sigma_{\eta\varepsilon} & 1 \end{bmatrix}\right)$, $\alpha_1 = 0.2$ and $\sigma_{\eta\varepsilon} = \sigma_{\varepsilon\eta} = 0.4$.

Figure 2: Finite Sample Power



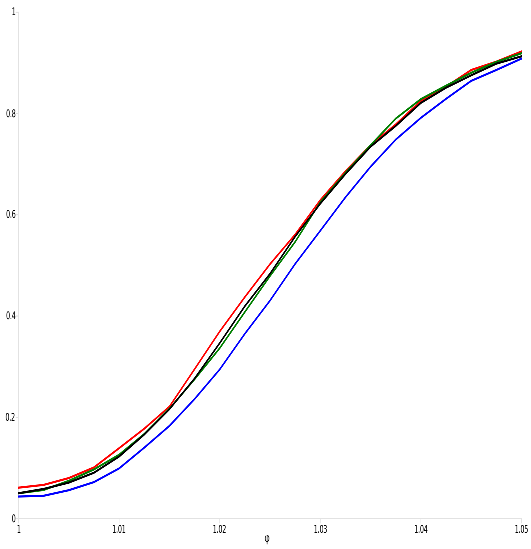
(a)

$\beta = 0.8 \quad \lambda = 0.5$



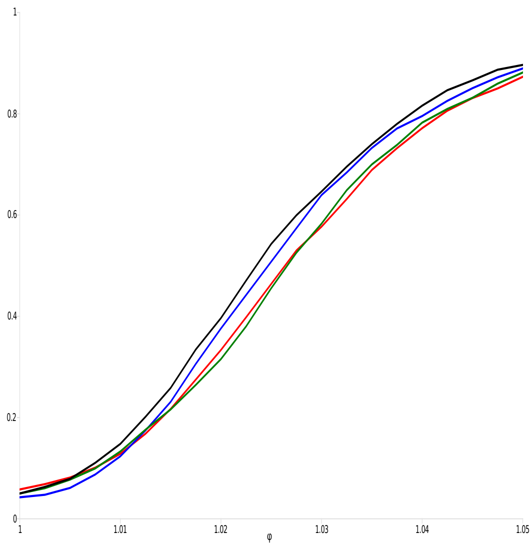
(b)

$\beta = 0.5 \quad \lambda = 0.5$



(c)

$\beta = -0.5 \quad \lambda = 0.5$



(d)

$\beta = -0.8 \quad \lambda = 0.5$

BSADF* test: —— CBSADF* test: ——
size-adjusted BSADF test: —— size-adjusted CBSADF test: ——

Data generated according to $y_t = y_{t-1} + u_t,\ t = 1, ..., 200$ and $y_t = \phi_t y_{t-1} + u_t, t = 201, ..., 250$ with $u_t = \alpha_1 u_{t-1} + v_t$, $v_t = \beta \Delta x_t + \varepsilon_t$ and $\Delta x_{t+1} = \lambda \Delta x_t + \eta_t$, where $\xi_t = \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{\varepsilon\eta} \\ \sigma_{\eta\varepsilon} & 1 \end{bmatrix}\right)$, $\alpha_1 = 0.2$ and $\sigma_{\eta\varepsilon} = \sigma_{\varepsilon\eta} = 0.4$.

3 Wild Bootstrap Testing for Speculative Bubbles Using Spot and Futures Prices

3.1 Introduction

Extant literature in economics and finance has recently expanded to include research work on financial bubbles. Although traditional analysis in econometrics did not allow for bubble episodes until recently, econometrics tests on bubble identification have emerged under the assumption of rational expectations, and their extensions can be used not only to identify historical episodes of explosive behaviour but real time detection of a bubble episode as well. According to the rational expectations theory, the real price of an asset should be equal to the present value of the future cash flows the asset generates, however this framework cannot explain deviations from equilibrium that are attributed to the existence of asset bubbles. This has led to the development of novel econometric methods that perform well under the rational expectations assumption.

3.1.1 Rational Bubble Tests

In their seminal study, Diba and Grossman (1988) put emphasis on the importance of unit root tests on asset bubble identification. Following the traditional approach of applying a standard left-tailed unit root to test the null hypothesis of a unit root to conclude that if both real prices and fundamentals (e.g. dividends) are non-stationary in levels but stationary in differences then that is indicative of non-existence of a rational bubble. The above approach has been criticised on its validity to detect periodically collapsing bubble episodes by Evans (1991) who argues that conventional unit root tests have low power on identifying bubble episodes due to their non-linear characteristics, therefore mean reversion when the asset bubble collapses may lead to such a price adjustment that the series might appear to have no explosive behaviour at all.

Rather than applying the left-tailed tests on the difference between real prices and fundamentals, recent research has put more emphasis on the use of right-tailed unit root tests that test the null hypothesis of a unit root against the alternative hypothesis of explosiveness. In particular, Phillips et al. (2011) propose a forward recursive right-tailed supremum ADF (SADF) test for rational bubbles that takes the maximum of a sequence of test statistics by changing the ending point of the subsamples while keeping the starting point fixed, and can identify bubble episodes that occur within the sample. Phillips et al. (2011) find evidence of the dot-com bubble in the beginning of the 2000's by applying the test on the NASDAQ stock price and dividend index. The Phillips et al. (2011) methodology offers the advantage that can be applied to a series of different assets such as commodities (see *inter alia* Gilbert 2010 and Homm and Breitung 2012) and exchange rates (Bettendorf and Chen 2013).

Homm and Breitung (2012) consider a modified version of the Phillips et al. (2011) SADF test together with a modified version of the locally best invariant (LBI) test of Busetti and Taylor (2004) and conclude that the SADF test has good power performance even when there are structural breaks in the sample and (or) the bubble episode occurs at the end of the sample. Empirically, they find statistical evidence of explosive behaviour before the 2008 global financial crisis in a number of countries including UK, US and Spain.

An important drawback of the SADF test of Phillips et al. (2011) is that it has low power in identifying multiple bubble episodes within the same sample leading to wrong statistical inference or pseudo-stationarity. For this reason, Phillips et al. (2015) suggest two recursive processes that are modifications of the SADF test of Phillips et al. (2011) that have great power on identifying periodically collapsing bubbles not only on historical data but in real-time as well. The two tests proposed by Phillips et al. (2015) are the the backward recursive SADF test (BSADF) that takes the maximum of a sequence of test statistics by changing the starting point of the subsamples while keeping the ending point fixed and a double-recursive generalised SADF test (GSADF) that takes the maximum of a sequence of test statistics by recursively changing the starting and ending points of the sample covering more subsamples of the data. Furthermore, the GSADF test seems to have better size and power performance compared to the SADF test and therefore can be utilised to identify multiple explosive episodes in a financial time series that can occur anywhere in the sample.

Although the GSADF test of Phillips et al. (2015) is rather powerful on identifying bubble episodes that occur in the sample, it cannot provide information on the origination and termination dates of the bubble episodes. For that reason, date-stamping techniques need to be applied by utilising the BSADF test which offers advantage on detecting end-of-sample bubble episodes as well. In their empirical application, Phillips et al. (2015) test for explosive behaviour in the S&P 500 stock price and dividend index from January 1871 to December 2010 by utilising the GSADF test and the BSADF test to date-stamp the bubble episodes to successfully identify historical banking crises and bubble episodes during this period.

In other empirical applications, Tsvetanov et al. (2016) apply the BSADF test to crude oil spot and futures contract prices over the period September 1995 to December 2013 to find that longer maturity contracts appear to indicate the origination date of the oil bubble episode of 2008 earlier compared to shorter maturity contracts, to conclude that futures contracts that have a maturity longer than six months have been significantly overpriced during that period.

More recently, Pavlidis et al. (2017) suggest that the relationship between future spot and futures contract prices can be disrupted as a result of periodically collapsing bubbles, market efficiency does not hold any more and therefore, futures contract prices are biased on predicting future spot prices with explosive degree of bias. Applying the Phillips et al. (2015) methodology to a series of different datasets, including the "German Hyperinflation" period (December 1921 to August 1923), the "Recent Float" period (January 1979 to December 2013) and the U.S equity market (December 1982 to March 2015) they argue that detecting explosivity in asset prices does not necessarily imply the existence of an asset price bubble in the series as explosive behaviour could also be present in fundamentals. Dealing with this inconclusive inference, Pavlidis et al. (2018) utilise market expectations, to study the crude oil market on the argument that market expectations are not influenced by the risk premium and they, once again, apply the Phillips et al. (2015) methodology to WTI crude oil spot and futures contract prices between January 1990 and December 2013. Finally, they conclude that the oil price run-up from 2004 to 2008 should only be attributed to the changes in fundamentals rather than a speculative bubble.

In this chapter we investigate whether the size and power properties of right-tailed Dickey-Fuller unit root test processes of Phillips et al. (2015) can be improved by applying a wild bootstrap approach that allows for potential heteroskedastic behaviour in the innovations that might be attributed to structural breaks, regime changes or volatility shifts to test for market efficiency in the commodity markets. For this reason, we model the series of interest as a moving average process rather than a unit root since under the null hypothesis of market efficiency, the difference between the future spot price and the futures contract price will be a stationary moving average process which order depends on the futures contract length.

We focus on the case that the explosive episode occurs at the end of the sample to identify these episodes in real time as this might be of importance to policy makers and central banks. The simulations results show that the proposed wild bootstrap test offers better size control and superior power performance in finite samples as the wild bootstrap test appears to be less size distorted compared to the non-bootstrap test whilst offering significant power gains. In the empirical application we apply the proposed and extant tests to the difference between the WTI crude oil future price and the price of nine futures contracts across different maturities over the period September 1995 to July 2019. We concentrate mainly on the 2007-2008 oil price run-up and the 2014-2015 oil price collapse and our proposed test identifies the two episodes while the conventional test of Phillips et al. (2015) does either not identify an episode at all, or identifies the origination date of the episode with delay reflecting the superior power of our proposed wild bootstrap test to effectively identify episodes of non-stationarity that occur at the end of the sample. Our proposed test suggests periods of market inefficiency prior to the existence of the bubble episode as identified by the conventional tests.

The remainder of this chapter is organised as follows. In Section 3.2 we outline the explosive financial bubble model and its assumptions, in Section 3.3 we review some extant tests whereas in Section 3.4 we present some limit theory. In Section 3.5 we introduce our proposed tests. In Section 3.6 the finite and sample size and power properties of our proposed tests are examined using Monte Carlo simulations. Section 3.7 presents an empirical application of our proposed tests to the WTI crude oil spot and futures contract prices. Section 3.8 concludes. Tables and Figures are presented in sections 3.9 and 3.10 respectively.

In the following $\xrightarrow{p}$ denotes convergence in probability, $\xrightarrow{d}$ denotes convergence in distribution, $\lfloor . \rfloor$ denotes the integer part of its argument and $y := x$ $(x := y)$ indicates that y is defined by x (x is defined by y).

3.2 The Model and Assumptions

Consider a rational expectations asset pricing model that relates the log of the spot price of oil, s_t to a fundamental component, v_t and a periodically collapsing speculative bubble b_t (see *inter alia* Sarno and Taylor 2003 and Engel and West 2005), such that;

$$s_t = v_t + b_t \tag{3.1}$$

as in Engel et al. (2007) where the fundamental price, v_t of an asset is equal to the stream of the discounted future cash flows generated by storing the commodity. In case of oil, this stream represents the convenience yield and refers to the process of storing inventories to meet unexpected changes in future market conditions.

To continue with, we let the fundamental v_t follow an autoregressive process of order one;

$$v_t = \phi v_{t-1} + \theta_t \tag{3.2}$$

where $\theta_t \sim^{iid} N(0, \sigma_\theta^2)$ a white noise process and $\phi \in \mathbb{R}$. In a rational bubble framework, following Blanchard (1979), we let b_t have two regimes that occur with probability π and $1 - \pi$ respectively. During the first regime, the bubble grows at a rate of $\frac{(1+r)}{\pi}$ exponentially, whereas in the second regime the bubble collapses to a white noise process;

$$b_{t+1} = \begin{cases} \left(\frac{1+r}{\pi}\right) b_t & \text{with probability } 1 - \pi \\ \varepsilon_{t+1}, & \text{with probability } \pi \end{cases} \tag{3.3}$$

where r is a constant discount rate and $\varepsilon_t \sim^{iid} N(0, \sigma_\varepsilon^2)$. Equation (3.3) is consistent with the rational bubble framework of Diba and Grossman (1988) and therefore;

$$E_t[b_{t+1}] = (1 + r)b_t, \tag{3.4}$$

where E_t is the expectation operator and again r is a constant discount rate that reflects the state of the economy. Under the assumption of risk neutrality, we can now define the logarithm of the price of the futures contracts of oil, $f_{t,n}$ with maturity n periods ahead as;

$$f_{t,n} = E_t[s_{t+n}] = E_t[v_{t+n}] + E_t[b_{t+n}]. \tag{3.5}$$

From Equations (3.2) and (3.4) we can now rewrite Equation (3.5) as;

$$f_{t,n} = \phi^n v_t + (1+r)^n b_t \tag{3.6}$$

and therefore from Equation (3.1), the future spot price n periods ahead is given by;

$$s_{t+n} = v_{t+n} + b_{t+n}. \tag{3.7}$$

Let the bubble component grow with probability π as in Equation (3.3) then substituting recursively;

$$s_{t+n} = \phi^n v_t + \left(\frac{1+r}{\pi}\right)^n b_t + \varepsilon^*_{t+n}, \tag{3.8}$$

where ε^*_{t+n} is a sum of two moving average (MA) processes;

$$\varepsilon^*_{t+n} = \sum_{i=1}^{n} \phi^{n-i}\theta_{t+i} + \sum_{i=1}^{n} \left(\frac{1+r}{\pi}\right)^{n-i} \varepsilon_i. \tag{3.9}$$

Comparing Equations (3.6) and (3.8) we can see that the future spot price is greater than the futures contract price or expected price as rational agents assign a non-zero probability to the bubble bursting and thus $(\frac{1+r}{\pi})^n$ is larger than $(1+r)^n$. Subtracting Equation (3.6) from Equation (3.8);

$$s_{t+n} - f_{t,n} = \left(\phi^n v_t + \left(\frac{1+r}{\pi}\right)^n b_t + \varepsilon^*_{t+n}\right) - \left(\phi^n v_t + (1+r)^n b_t\right) = (1+r)^n \left(\frac{1}{\pi^n} - 1\right) b_t + \varepsilon^*_{t+n}. \tag{3.10}$$

As can be seen in Equation (3.10), $s_{t+n} - f_{t,n}$ is a linear function of two moving average (MA) processes and a bubble process and therefore exhibits explosive behaviour. From Equation (3.10), it is evident that $s_{t+n} - f_{t,n}$ does not depend on market fundamentals and thus any evidence of explosiveness can only be attributed to future spot or futures contract prices.

3.3 Extant Recursive Test Procedures

Due to the poor power performance of the standard univariate ADF test on detecting short lived explosive episodes in full samples, Phillips et al. (2015) introduce a univariate approach to testing for bubble episodes that occur at the end of the sample by utilising the standard ADF regression;

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{k=1}^{p} \alpha_k \Delta y_{t-k} + e_t \tag{3.11}$$

performed on subsamples of the data, where μ is the intercept and p is the number of lags of the dependent variable Δy_t and the test statistics are function of a sequence of ADF statistics of the subsamples. In particular, if we denote the ADF test applied to the full sample as $ADF_0^1(p)$, then the ADF test procedure performed on the subsample $t = \lfloor r_1 T \rfloor, ..., \lfloor r_2 T \rfloor$ can be denoted as $ADF(p)_{r_1}^{r_2}$. Phillips et al. (2015) propose the following test statistic to test for an explosive episode;

$$SADF := \sup_{r_2 \in [r_0, 1]} \{ADF(p)_0^{r_2}\}. \tag{3.12}$$

Subject to a minimum sample size $\lfloor r_0 T \rfloor$, the SADF test is the supremum of right-tailed ADF statistics performed on all subsamples starting at $t = 1$. Although the SADF test is powerful on detecting periodically collapsing bubble episodes and it can be very useful in the construction of confidence intervals, the performance of the test on detecting end-of-sample explosive episodes, that can be of interest to policy makers and regulators, is rather low as the test uses more observations at the beginning rather than the end of the sample.

Focusing on explosive episodes that occur at the end of the sample and motivated by the low power to detect these episodes Phillips et al. (2015) propose utilising the following test statistic instead;

$$BSADF := \sup_{r_1 \in [0, 1-r_0]} \{ADF(p)_{r_1}^{1}\}. \tag{3.13}$$

Subject to a minimum sample size $\lfloor r_0 T \rfloor$, the BSADF test is the supremum of right-tailed ADF statistics computed on all subsamples ending at date $t = T$. Using more observations at the end of the sample, the BSADF test is designed in such a way that it is particularly powerful when the explosive episode occurs at the end of the sample and can be rather useful for date stamping past bubble episodes.

Finally, Phillips et al. (2015) propose the GSADF test that is a double-recursive unit root test as well, constructed from a sequence of ADF test statistics computed over all possible start and end dates of the subsamples, subject to a minimum sample size $\lfloor r_0 T \rfloor$;

$$GSADF := \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \{ADF(p)_{r_1}^{r_2}\}. \tag{3.14}$$

The GSADF test is designed to test for the presence of one or multiple bubble episodes in a financial time series that can occur anywhere in the sample. Therefore, when using the GSADF test a rejection of the null in favour of explosive behaviour can only indicate the existence of a bubble episode but not exactly where in the sample the episode occurs. Date stamping techniques based on the BSADF test might be utilised instead.

3.4 Limit Theory

In this section we present the limiting null distribution of the GSADF test statistic, with the limiting null distributions of the SADF and BSADF test statistics following as special cases of the GSADF one.

The limiting null distribution of the GSADF test statistic is outlined in the following theorem and applies only to the case that the series is a unit root process under the null hypothesis.

Theorem 3.3. *When the regression model includes an intercept and under the null hypothesis has a unit root then as in Phillips et al. (2015);*

$$GSADF \xrightarrow{d} \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \left\{ \frac{\frac{1}{2} r_w \left[W(r_2)^2 - W(r_1)^2 - r_w \right] - \int_{r_1}^{r_2} W(r) dr \left[W(r_2) - W(r_1) \right]}{r_w^{1/2} \left(r_w \int_{r_1}^{r_2} W(r)^2 dr - \left[\int_{r_1}^{r_2} W(r) dr \right]^2 \right)^{1/2}} \right\} \tag{3.15}$$

where $r_w = r_2 - r_1$ and W is a standard Brownian motion process.

Remark 3.7. *The asymptotic distribution of the BSADF test is a special case of the above where* $r_w = 1 - r_1$, $r_1 \in [0, 1 - r_0]$ *and* $r_2 = 1$ *whereas the asymptotic distribution of the SADF test can be found by setting* $r_w = r_2$, $r_2 \in [0, 1]$ *and* $r_1 = 0$.

3.5 Proposed Tests

The limit theory presented above only holds under the null hypothesis of a unit root, therefore applying any of the extant recursive tests of Phillips et al. (2015) on the difference between the future spot price and the futures contract price with maturity n will be inaccurate and will lead to wrong statistical inference since under the null hypothesis of no bubble behaviour, the series is a moving average process $MA(n)$ where n is the length of the futures contract.

In particular, consider a rational expectations asset pricing model that relates the log of spot price of oil, s_t to a fundamental component, v_t under the null hypothesis of no bubble episode;

$$s_t = v_t \tag{3.16}$$

where again, the fundamental price, v_t of an asset is equal to the stream of the discounted future cash flows generated by storing the commodity. As previously, we let the fundamental v_t follow an autoregressive (AR) process of order one as in Equation (3.2).

In the absence of speculative bubbles and under the assumption of risk neutrality, the logarithm of the price of the futures contract of oil, $f_{t,n}$ with maturity n periods ahead is given by;

$$f_{t,n} = E_t[s_{t+n}] = E_t[v_{t+n}]. \tag{3.17}$$

And therefore from Equation (3.2) we can rewrite Equation (3.17) as;

$$f_{t,n} = \phi^n v_t \tag{3.18}$$

and then from Equation (3.16), the future spot price n periods ahead will be given by;

$$s_{t+n} = v_{t+n}. \tag{3.19}$$

Substituting recursively from Equation (3.2) will give;

$$s_{t+n} = \phi^n v_t + \varepsilon^*_{t+n}, \tag{3.20}$$

where ε^*_{t+n} is a sum of a moving average (MA) process as;

$$\varepsilon^*_{t+n} = \sum_{i=1}^{n} \phi^{n-i} \theta_{t+i}. \tag{3.21}$$

A comparison between Equations (3.18) and (3.20) shows that the future spot price is greater than the futures contract price or expected price by the error term ε^*_{t+n}. Subtracting Equation (3.18) from Equation (3.20);

$$s_{t+n} - f_{t,n} = \left(\phi^n v_t + \varepsilon^*_{t+n}\right) - \left(\phi^n v_t\right) = \varepsilon^*_{t+n}. \tag{3.22}$$

Under the null hypothesis of no bubble, Equation (3.22) $s_{t+n} - f_{t,n}$ is a stationary moving average process. As we will see in the next section, rejection of the null hypothesis indicates the existence of explosive behaviour in the sample.

Applying any of the extant recursive tests of Phillips et al. (2015) on the difference of the series generated by Equation (3.22) will, therefore, lead to downward size distortions as $s_{t+n} - f_{t,n}$ is a stationary moving average process under the null hypothesis, rather than a unit root process. In other words, under the null hypothesis of no bubble behaviour, the series is a moving average process $MA(n)$ where n is the length of the futures contract.

Given that the critical values are computed under the null hypothesis of a unit root, we expect that the extant recursive tests of Phillips et al. (2015) will be severely undersized. Thus, we impose a moving average behaviour in the regression model and we propose utilising the following wild bootstrap implementation of the Phillips et al. (2015) tests that leads to improved size control, while offering significant power gains.

3.5.1 Wild Bootstrap Unit Root Tests

To better control size in finite samples we follow an approach that applies the wild bootstrap algorithm presented below to the first differences of our series, constructing a wild bootstrap version of the BSADF test of Phillips et al. (2015) that can control the size and power performance of the test. As can been seen in Equation (3.22), under the null hypothesis of stationarity, and thus in absence of a bubble, the first difference between future spot prices and futures contract prices is equal to $\{\varepsilon^*_{t+n}\}$ which is a moving average $MA(n)$ process, where the order n is determined by the futures contract length.

We choose to apply the wild bootstrap approach over the i.i.d. bootstrap as the former allows for potential heteroskedastic behaviour in the innovations. Similarly to Harvey et al. (2017) we utilise the following bootstrap algorithm.

Algorithm 2. ***(Wild Bootstrap GSADF test)***

Step 1: *Set* $\theta_t = \Delta s_t$ *where* Δ *is the first difference operator and t=1,...,T.*

Step 2: *Construct* $\theta^*_t = w_t\theta_t$ *where* $\{w_t\}^T_{t=1}$ *and* $w_t \sim^{iid} N(0,1)$ *a random sequence.*

Step 3: *Construct* $\{\varepsilon^*_t\}$ *the bootstrap sample as a partial sum process;*

$$\varepsilon^*_t := \sum_{j=1}^{n} \theta^*_{t-j+1}, \text{ for } t = 1, ..., T \tag{3.23}$$

where n is the periods to maturity.

Step 4: *Set* $y^*_t = \varepsilon^*_t$*, t=1,...,T.*

Step 5: *Compute the bootstrap test statistic for the GSADF test;*

$$GSADF^*_{wb} := \sup_{\substack{r_2\in[r_0,1] \\ r_1\in[0,r_2-r_0]}} \{ADF^{*r_2}_{r_1}\}. \tag{3.24}$$

where $ADF^{*r_2}_{r_1}$ *is the test statistic of* $\hat{\phi}^*_{r_1,r_2}$ *in the fitted OLS regression;*

$$\Delta y^*_t = \hat{\alpha}^*_{r_1,r_2}\hat{\phi}^*_{r_1,r_2}y^*_{t-1} + u^*_t \tag{3.25}$$

calculated in a double-recursive framework over all possible start (r_1) and end (r_2) points of the sample. Therefore;

$$ADF^{*r_2}_{r_1} = \frac{\hat{\phi}^*_{r_1,r_2}}{s(\hat{\phi}^*_{r_1,r_2})} \tag{3.26}$$

where the standard errors $s(\hat{\phi}^*_{r_1,r_2})$ *are defined by;*

$$s^2(\hat{\phi}^*_{r_1,r_2}) = \frac{\hat{\sigma}^{*2}_{r_1,r_2}}{\Sigma' y^{*2}_{t-1} - \left(\Sigma' y^*_{t-1}\right)^2 / \Sigma' 1}, \tag{3.27}$$

and $\hat{\sigma}^{*2}_{r_1,r_2} = \left(\frac{1}{[Tr_w]}\right)\Sigma' \hat{\varepsilon}^*_t$. *The BSADF test is a special case of GSADF test where* $r_w = 1 - r_1$, $r_1 \in [0, 1 - r_0]$ *and* $r_2 = 1$ *whereas the SADF test can be defined by setting* $r_w = r_2$, $r_2 \in [0, 1]$ *and* $r_1 = 0$.

Remark 3.8. *The algorithm is outlined for the case where no deterministic components are allowed for in the data. If a constant is to be allowed for in the data then the series* y_t *should be replaced by their demeaned counterparts. Likewise, if a constant and trend are to be allowed for in the data then* y_t *should first be demeaned and detrended.*

The bootstrap errors $\{\varepsilon^*_t\}$ can, potentially, replicate the pattern of heteroskedasticity as $\varepsilon^*_t \sim^{iid} N(0, (\Delta y_t)^2)$ conditional on Δy_t. We suggest utilising the ADF^* test statistic in place of the standard ADF test statistic to deliver a test with controlled finite sample size when estimation is performed in a recursive framework. The wild bootstrap BSADF test is a special case of the wild bootstrap GSADF test as defined in Equation (3.24) where $r_w = 1 - r_1$, $r_1 \in [0, 1 - r_0]$ and $r_2 = 1$ whereas the SADF test can be defined by setting $r_w = r_2$, $r_2 \in [0, 1]$ and $r_1 = 0$. The wild bootstrap BSADF and SADF tests are defined as;

$$BSADF^*_{wb} := \sup_{r_1 \in [0, 1-r_0]} \{ADF^1_{r_1}\} \tag{3.28}$$

$$SADF^*_{wb} := \sup_{r_2 \in [r_0, 1]} \{ADF^{r_2}_0\}. \tag{3.29}$$

The benefit of the wild bootstrap procedure is that the critical values of the recursive bootstrap ADF tests, rather than assuming a unit root process under the null hypothesis like the conventional recursive ADF tests do, are obtained by imposing a $MA(n)$ process for our series under the null of the absence of speculative bubbles in the sample. The wild bootstrap ADF tesst will be shown to better control size in finite samples compared to the non-bootstrap version of the tests.

In what follows we will focus on the performance of the BSADF$^*_{wb}$ test as it is constructed to detect end-of-sample and ongoing explosive episodes which are arguably of most interest empirically.

3.6 Finite Sample Simulations

In this section we will examine the finite sample size and power performance of our proposed test relative to the extant tests of Phillips et al. (2015). In order to do so, data were simulated according to the following data generating process;

$$y_{t+n} = s_{t+n} - f_{t,n} \quad \text{for } t = 1, ..., T. \tag{3.30}$$

From Equation (3.10) we know that;

$$s_{t+n} - f_{t,n} = b_t + \varepsilon^*_{t+n} \quad \text{for } t = 1, ..., T \tag{3.31}$$

and therefore;

$$y_{t+n} = b_t + \varepsilon^*_{t+n} \quad \text{for } t = 1, ..., T \tag{3.32}$$

where;

$$\varepsilon^*_{t+n} = \sum_{i=1}^{n} \phi^{n-i} \theta_{t+i} \tag{3.33}$$

and where $\theta_t \sim^{iid} N(0, \sigma^2_\theta)$ a white noise process, n is the length of the futures contract and;

$$b_t = \begin{cases} 0, \text{ for } t = 1, ..., t_b \\ \phi_t b_{t-1} + \varepsilon_t, \text{ for } t = t_b + 1, ..., T \end{cases} \tag{3.34}$$

an end-of-sample bubble process where $\varepsilon_t \sim^{iid} N(0, \sigma^2_\varepsilon)$.

Under the null hypothesis of no speculative bubble, the series of interest $\{y_{t+n}\}$ is a moving average $MA(n)$ process where n is the length of the futures contract;

$$y_{t+n} = \varepsilon^*_{t+n} \quad \text{for } t = 1, ..., T \tag{3.35}$$

whereas, under the alternative hypothesis of an end-of-sample bubble the series of interest $\{y_{t+n}\}$ is a linear function of a moving average (MA) process just as in Equation

(3.35) for $t = 1, ..., t_b$ where t_b is the origination point of the speculative bubble, whereas for $t = t_b + 1, ..., T$ the series y_{t+n} is sum of two moving average (MA) processes plus a bubble component;

$$y_{t+n} = \begin{cases} \varepsilon^*_{t+n} & \text{for } t = 1, ..., t_b \\ b_t + \varepsilon_{t+n} + \varepsilon^*_{t+n}, & \text{for } t = t_b + 1, ..., T. \end{cases} \tag{3.36}$$

We use a fixed lag order of zero for the dependent variable Δy_t as in Harvey et al. (2017) as the wild bootstrap resampling introduced in Step 2 of Algorithm 2 wipes out any weak dependence present in Δy_t. This implies that there is no certain requirement to augment the sub-sample regressions underlying the wild bootstrap procedure of the BSADF test with lagged-difference regressors. In fact, as Harvey et al. (2017) point out, the wild bootstrap BSADF test is both asymptotically consistent and consistent against fixed magnitude bubble alternatives for any lag length of Δy_t.

The minimum window size, r_0 for all recursive test procedures is chosen as;

$$r_0 = (0.01 + \frac{1.8}{\sqrt{T}}) * T. \tag{3.37}$$

All simulations that follow were conducted in GAUSS 17 using $10,000$ Monte Carlo replications and 399 bootstrap replications. All tests are performed at a nominal 5% level of significance. The sample size is set equal to $T = 200, 400$ and 800 and $\sigma_\varepsilon = 0.3$, $\sigma_{\varepsilon^*} = \sqrt{0.1574}$ and $\sigma_\theta = 1$ as in Pavlidis et al.(2017).

3.6.1 Empirical Size

To assess the size performance of our proposed test, data were generated according to Equation (3.35) for the full sample $t = 1, ..., T$. We report the empirical size of both our proposed BSADF$^*_{wb}$ test and the BSADF test of Phillips et al. (2015) for a range of different lengths of the futures contracts, n.

Table 4 reports the finite sample size of all tests for sample sizes of $T = 200, 400$ and 800. When $T = 200$, the non-bootstrap BSADF test exhibits severe size distortions, with the size of this test far below the nominal 5% level of significance across all contracts. The poor size control of the BSADF test should be attributed to the stationary behaviour of the series of interest y_{t+n} under the null hypothesis since that series, in the absence of a bubble, is equal to a moving average process $MA(n)$ where the order n is determined by the futures contract length. The limit null distribution and critical values of the BSADF test are a generated under the assumption of a unit root and, therefore, these incorrect critical values result in severe size distortions.

The wild bootstrap based BSADF$^*_{wb}$ test displays much better size control overall. In particular, the BSADF$^*_{wb}$ statistic appears to be slightly oversized across different contract lengths exhibiting size ranging from 0.054 to 0.075, while in most cases the BSADF$^*_{wb}$ statistic has reasonably controlled size across most of the contract lengths. Some modest oversize that is exhibited for the three month and six month futures contract could just be considered a small sample issue that we can easily account for by increasing the sample size.

As can been seen in Table 4, when we increase the sample size to 400 and 800, results for the non-bootstrap based BSADF test are broadly similar to those reported for $T = 200$, with these tests still displaying even more undersizing across all futures contracts, indicating that the size distortions exhibited by this test is not simply a small sample issue. The BSADF$^*_{wb}$ statistic appears to be less oversized across all contracts and the size of the BSADF$^*_{wb}$ test is still reasonably well controlled overall, although this test still exhibits some modest oversize for the three month futures contract similarly to sample size of $T = 200$, when the sample size is $T = 400$ whereas no significant over-sizing is observed when the sample size increases to $T = 800$. The BSADF$^*_{wb}$ tests displays good size control across all different contract lengths, with the finite size values greatly improved in the larger sample size of $T = 400$.

Overall, it can be seen that the bootstrap BSADF$^*_{wb}$ test controls size to a much greater degree than the respective non-bootstrap BSADF test, with the latter displaying such severe undersize that it would be of little use empirically in the scenarios considered. Whilst the wild bootstrap BSADF$^*_{wb}$ test displays better size control than its non-bootstrap counterpart, it still displays some modest oversize across a number of contracts with different length. Overall, the BSADF$^*_{wb}$ test, has the best overall size control, with the modest oversize exhibited by this test in smaller sample sizes almost entirely eliminated when a larger sample size is considered.

3.6.2 Empirical Power

We now proceed to examine the power performance of our proposed tests relative to extant tests. To do so, data were generated according to Equations (3.34) and (3.36) with $T = 200$ under the alternative hypothesis of and end-of-sample bubble episode by setting $b_t = 0$ for t=1,...,180, and $\phi_t = \phi > 1$ for t=181,...,200. The series $\{y_t\}$ follows a stationary process for the first 180 observations and is then subject to (potential) explosive behaviour over the remaining 20 observations. In what follows we will focus on the power performance of the BSADF and BSADF$^*_{wb}$ tests as they are constructed to detect end-of-sample and ongoing explosive episodes which are arguably of most interest empirically.

The finite sample power of the BSADF and BSADF$^*_{wb}$ tests was computed for a grid of 50 values of ϕ from $\phi = 1.00$ to $\phi = 1.30$ for each of the nine different futures contract lengths previously considered. Figures 3, 4 and 5 report finite sample power curves for the BSADF and BSADF$^*_{wb}$ tests across each of the contract lengths considered. In all different contract lengths the power of both tests is increasing monotonically in ϕ, although for short-mid length futures contracts [Figure 3, (a) to (d)] the power of the BSADF$^*_{wb}$ test surges for lower values of ϕ while the BSADF test appears to have zero power due to the undersize exhibited by the BSADF test across all contract lengths. In general, the power of the BSADF$^*_{wb}$ test exceeds that of the BSADF test for different values of ϕ and contract lengths.

The BSADF$^*_{wb}$ test displays much greater power than the BSADF test, with the power differential between the two tests reaching almost 70% for low values of ϕ as can be seen in Figure 3 (a). In Figures 4 (c), (d) and 5 (a) the BSADF$^*_{wb}$ test offers relatively little additional power compared to the BSADF test, however we may argue that this could be reasonably attributed to the fact that for long length futures contracts, the series of interest y_{t+n} becomes more persistent under the null hypothesis as the contract length increases and, therefore, the wild bootstrap simulated critical values are closer to the critical values simulated under the null hypothesis of a unit root. Both the BSADF test and the BSADF$^*_{wb}$ appear to have similar power performance for very large values of ϕ for mid and long length contracts [Figures 4 (a)-(d) and 5 (a)].

Overall, we argue that the bootstrap BSADF$^*_{wb}$ test shows better size control than its respective non-bootstrap BSADF test across all futures contract lengths. Arguably, applying the wild bootstrap procedure on the BSADF test leads to greater power performance relative to the non-bootstrap BSADF test in finite samples as well as offering improved size control.

The benefit of the wild bootstrap procedure is that the critical values of the recursive bootstrap ADF tests do not consider a unit root process under the null hypothesis but resemble the behaviour of a moving average process MA(n) instead, where the order n is determined by the length of the futures contract as in Equation (3.22), in the absence of speculative bubbles in the sample. In fact, the wild bootstrap ADF test is shown to better control size in finite samples compared to the non-bootstrap version of the tests. We, therefore, recommend utilising the BSADF^*_{wb} test in practice as it offers the best overall size control and power properties between the tests considered.

3.7 Empirical Application

To demonstrate the usefulness of our proposed test we consider the following empirical application. We download WTI crude oil spot and futures contract prices from Eikon for the period September 1995 to July 2019 at weekly and monthly frequency. The futures contract maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months as in Tsvetanov et al. (2016) and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month.

Next, we apply the BSADF and BSADF^*_{wb} tests as defined by Equations (3.13) and (3.29) respectively. Furthermore, in computing the tests-statistics and following Harvey et al. (2017) for the standard BSADF test we allow for a maximum six lags of the differenced dependent variable to account for serial correlation and we let the Bayesian Information Criterion (BIC) decide on the optimal lag structure whereas for simulating the critical values for the wild bootstrap BSADF^*_{wb} test we include no lagged difference augmentation since as explained earlier, the wild bootstrap resampling introduced in Step 2 of Algorithm 2 annihilates any weak dependence present in Δy_t and therefore there is no certain requirement to augment the sub-sample regressions underlying the wild bootstrap procedure of the BSADF test with lagged-difference regressors.

To continue with, we compute right-tailed finite sample critical values for both tests using $10,000$ Monte Carlo replications for the BSADF test and $9,999$ bootstrap replications for the BSADF$^*_{wb}$ test respectively. The minimum window size is determined as in Equation (3.37) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regression.

Following Phillips et al. (2015), who perform a (pseudo) real-time bubble monitoring exercise on the present value of the real S&P 500 price-dividend ratio and apply a date-stamping strategy to test for the presence of explosive behaviour, we investigate the power of our proposed tests to detect bubble episodes in commodity price series. In particular, we estimate both the BSADF and BSADF$^*_{wb}$ test statistics in a recursive framework on the difference between future spot and futures contract prices, $s_{t+n} - f_{t,n}$ where n represents the length of the contract ($n = 1, 3, 6, 9, 12, 15, 18, 21$ and 24), for all nine contracts. We define the origination date of the bubble episode as the first chronological observation of which the test statistic is larger than the simulated finite sample critical value therefore rejecting the null hypothesis of stationarity. Similarly, the termination date of the bubble episode is defined as the first chronological observation of which the test-statistic becomes smaller than the simulated finite sample critical value following the commencement of a bubble episode.

The proposed wild bootstrap test identifies periods of market inefficiency, that existent bubble tests do not detect, due to the imprecise estimation of the critical values as well as bubble episodes since under the alternative hypothesis, market inefficiency can be either attributed to a unit root or an explosive episode in the series and therefore can be used in practise as an early warning mechanism of market inefficiency that could, potentially, result in a bubble episode.

Interested in examining explosive episodes that occur at the end of the sample, we focus on the performance of the BSADF and BSADF$^*_{wb}$ tests as identifying bubble episodes in real-time can be useful to regulators, policy makers and central banks. Figures 6 to 10 and 11 to 15 plot the WTI crude oil future spot and futures contract logarithmic prices together with their difference across different contract lengths as well as the recursive BSADF test statistic against the corresponding 95% simulated Monte Carlo and wild bootstrap critical value sequences across all different lengths of the futures contracts both at weekly and monthly frequency respectively.

Tables 5 and 6 present the origination and termination dates of two explosive episodes, namely the 2007-2008 oil price run-up and the 2014-2015 oil price collapse, at weekly and monthly frequency respectively, as identified by the Phillips et al. (2015) BSADF test and our proposed BSADF^*_{wb} test across different futures contract lengths.

When data is used at weekly frequency, the BSADF test does not detect non-stationary behaviour for short maturity futures contracts up to three months whereas the BSADF^*_{wb} test suggests large periods of non-stationarity for short and mid maturity (up to six months) futures contracts. The proposed test seems to not be able to make a statistical inference on which part of this large period of non-stationarity can be attributed to a unit root and which to a, potential, bubble episode resulting in market inefficiency (see Table 5). However, for longer maturity futures contracts up to eighteen months, the BSADF test identifies two main explosive episodes, one that starts between April 2008, for the fifteen month contract and December 2008 for the twelve month contract and one that starts in November-December 2014, as can been seen in Table 5. For the same maturity futures contracts, our proposed BSADF^*_{wb} test identifies a period of market inefficiency prior to the 2007-2008 oil price run-up which starts between September 2006 for the eighteen month contract and December 2007 for the fifteen month contract.

In Table 6 we do the date stamping using data at monthly frequency and we see that for most of maturities of the futures contracts the BSADF test does not identify any bubble episodes at all except for the six month and twenty one month futures contract as in Figures 6 a) and 14 b) respectively. In particular, the proposed BSADF^*_{wb} test seems to be able to identify multiple episodes on non-stationarity that can be attributed to either market inefficiency or a bubble episode across all nine futures contracts with the proposed BSADF^*_{wb} test identifying both the WTI crude oil price run-up of 2007-2008 and the crude oil price collapse of 2014-2015.

Looking at the difference between the one month crude oil future spot price and the one month futures contract price at monthly frequency, $s_{t+1} - f_{t,1}$ presented in Table 6, we can see that for short-period maturity contracts the BSADF^*_{wb} test identifies large periods of non-stationarity without being able to distinguish between a unit root and an explosive episode that results in market inefficiency. For mid and longer maturity contracts, namely for the six month futures contract the BSADF^*_{wb} test identifies the origination episode of the 2014-2015 oil price collapse two months earlier than the BSADF test whereas for the twenty one month contract the proposed BSADF^*_{wb}

test identifies the origination date of the 2007-2008 episode in June 2007 similar to the BSADF test.

Utilising our proposed BSADF^*_{wb} test, on the difference between the three month crude oil future spot price and the three month futures contract price, $s_{t+3} - f_{t,3}$, we identify two episodes that suggest market inefficiency that can be either attributed to a unit root or an explosive episode, as seen in Table 6. The first episode occurs during the period March 2007 to January 2009 and the second one between January 2015 and July 2015. On the difference between the six month crude oil future spot price and the six month futures contract price, $s_{t+6} - f_{t,6}$, Phillips et al. (2015) BSADF test identifies the origination date of 2014-2015 episode as June 2015 whereas the wild bootstrap BSADF^*_{wb} test identifies the origination date two months earlier in April 2015 (Table 6).

The origination dates of the 2007-2008 and 2014-2015 episodes are consistent across mid and long maturity contracts as estimated by the BSADF^*_{wb} test at monthly frequency. Particularly, for the twelve month, fifteen month and eighteen month futures contract, the oil price episode of 2007-2008 is originated in June/July 2007 whereas the 2014-2015 collapse episode is originated in September/November 2015. When applied on the difference between the twenty one month crude oil future spot price and the twenty one month futures contract price, $s_{t+21} - f_{t,21}$, the BSADF test suggests June 2007 as the origination date of the 2007-2008 episode the same as the BSADF^*_{wb} test as can be seen in Table 6.

Figures 6 to 10 and 11 to 15 plot the series of interest together with the test-statistics and critical values at weekly and monthly frequency. On the left-hand side we can see the WTI crude oil future spot and futures contract prices in logarithms from September 1995 to July 2019 together with the difference between the future spot and futures contract prices in logarithms. On the right-hand side we see the BSADF test-statistics sequence together with the simulated BSADF and BSADF^*_{wb} finite sample critical values. In all Figures 6 to 15 we can see the series that represents the difference between the crude oil future spot and futures contract logarithmic prices. The series seem to exhibit periods of non-stationarity especially for mid and longer maturity contracts (see for example Figures 7 to 10 and 13 to 15) which resemble characteristics of a unit root.

3.7.1 Testing for Autocorrelation

Sample autocorrelation functions are presented in Figures 16 to 28 and 19 to 21 when data is at weekly and monthly frequency respectively. In both cases, non-stationary behaviour looks apparent. At weekly frequency, the error term seems to have a highly persistent effect on the current value of the $s_{t+n}-f_{t,n}$ series especially for longer maturity futures contracts. As can be seen in Figure 16 a), shocks to the difference between the one month crude oil future spot price and the one month futures contract price, $s_{t+1} - f_{t,1}$, die away relatively faster compared to longer maturities, however there is some persistence in the system that only goes away after eleven lags. Autocorrelation appears to be larger for longer maturity futures contracts as the sample autocorrelation function remains close to unity in some cases, suggesting that shocks are highly persistent and might remain in the system indefinitely especially for long maturity contracts (see Figure 17 d) and 18 a)).

Similar pattern is observed at monthly frequency. Autocorrelation appears to grow with the maturity of the futures contracts indicating persistent shocks in the system. At monthly frequency, shocks to the difference between the one month crude oil future spot price and the one month futures contract price, $s_{t+1} - f_{t,1}$, gradually decay after five lags (see Figure 19 a)), whereas shocks to the difference between the three (six) month crude oil future spot price and the three (six) month futures contract price, $s_{t+3} - f_{t,3}$ ($s_{t+6}-f_{t,6}$) fade out after four (two) lags therefore persistence is relatively low. For longer maturity contracts, autocorrelation is still persistent although it gradually decreases (see Figure 20 d) and 21 a)).

High autocorrelation suggests that shocks in the system are rather persistent suggesting that the series exhibit non-stationary behaviour that can be attributed to either a unit root or an explosive episode. At weekly frequency, the BSADF test and the BSADF$^*_{wb}$ test seem to agree on the origination dates of the two oil price episodes for mid and longer maturity contracts and the proposed BSADF$^*_{wb}$ test identifies the origination dates a few months earlier than the BSADF test for some futures contracts. In addition, prior to the 2007-2008 oil price run-up the BSADF$^*_{wb}$ test is able to identify periods of market inefficiency that can be related to either a unit root or explosive behaviour of the series. At monthly frequency, our proposed BSADF$^*_{wb}$ test seems to be able to identify more periods of non-stationarity compared to the BSADF test seems to be able to identify periods of market inefficiency that could either be attributed to a unit root or an explosive episode. When both tests identify such periods, the BSADF^*_{wb} test identifies the episode earlier for some futures contracts.

Overall, our proposed BSADF^*_{wb} test has identified the WTI crude oil price run-up of 2007-2008 and the WTI crude oil price collapse of 2014-2015 earlier than the BSADF test across futures contracts of different maturities whereas for some futures contracts the BSADF test does not identify any non-stationarity at all, while at the same time our proposed $BSADF_{wb}$ test indicates periods of market inefficiency prior to these episodes, reflecting the superior power of the BSADF^*_{wb} test to quickly detect non-stationary episodes that can be attributed to market inefficiency.

Our empirical result is consistent with the theoretical evidence presented in section 3.6 where we suggest that the wild bootstrap BSADF^*_{wb} test shows better size and power properties compared to the BSADF test and in this empirical exercise we present empirical evidence that utilising the wild bootstrap version of the BSADF test of Phillips et al. (2015) in finite samples results in improved power performance and therefore, advise utilising the BSADF^*_{wb} test in practise.

3.8 Conclusion

This chapter examines whether the size and power performance of right-tailed Dickey-Fuller unit root tests can be improved by applying a wild bootstrap approach to Phillips et al. (2015) tests to account for potential heteroskedasticity that might be attributed to structural breaks, regime changes or volatility breaks and deal with the size distortions of the BSADF test when applied on a series that replicates the difference between future spot and futures contract prices as in Pavlidis et al. (2018) to test for market efficiency in the commodity markets. Mainly interested in identifying explosive episodes in real-time, we focus on end-of-sample bubble episodes. The simulations results show that the proposed wild bootstrap test offers better size control and power performance in finite samples. Particularly, the wild bootstrap test appears to be less size distorted compared to the non-bootstrap test while the power gains are significantly higher.

In the empirical exercise we apply the proposed and extant tests on the difference between the WTI crude oil future spot price and the price of nine futures contracts across different maturities over the period September 1995 to July 2019. Focusing mainly on the 2007-2008 oil price run-up and the 2014-2015 oil price collapse, our proposed test identifies the two episodes earlier as the wild bootstrap BSADF^*_{wb} test suggests periods of non-stationarity that indicate market inefficiency prior to the 2007-2008 oil price run-up while the conventional test of Phillips et al. (2015) under performs our proposed wild bootstrap BSADF^*_{wb} test.

In summary, the wild bootstrap BSADF^*_{wb} test shows better size control than their corresponding non-bootstrap BSADF test across different maturity contracts, while the BSADF^*_{wb} test leads to greater power performance relative to the BSADF test in finite samples and offers significantly improved size control. Our proposed BSADF^*_{wb} test identifies the 2007-2008 oil price run-up and the 2014-2015 oil price collapse when the BSADF test does either not identify any explosive episode at all or identifies the explosive episode with delay, reflecting the superior power of the BSADF^*_{wb} test to identify episodes of non-stationarity (unit root or explosive) that can be attributed to market inefficiency.

3.9 Tables

Table 4: Finite Sample Size

	$T = 200$	
Contract Length (Months)	$BSADF$	$BSADF_{wb}^{*}$
1	0.000	0.069
3	0.000	0.075
6	0.001	0.064
9	0.007	0.068
12	0.011	0.060
15	0.017	0.054
18	0.024	0.058
21	0.028	0.058
24	0.030	0.055
	$T = 400$	
Contract Length (Months)	$BSADF$	$BSADF_{wb}^{*}$
1	0.000	0.063
3	0.000	0.067
6	0.000	0.063
9	0.004	0.062
12	0.001	0.061
15	0.003	0.059
18	0.004	0.056
21	0.008	0.057
24	0.010	0.057
	$T = 800$	
Contract Length (Months)	$BSADF$	$BSADF_{wb}^{*}$
1	0.000	0.059
3	0.000	0.059
6	0.000	0.060
9	0.000	0.063
12	0.000	0.055
15	0.000	0.061
18	0.000	0.055
21	0.000	0.053
24	0.001	0.054

Data generated according to $y_{t+n} = s_{t+n} - f_{t,n}$ with $s_{t+n} - f_{t,n} = \varepsilon_{t+n}^{*}$, $\varepsilon_{t+n}^{*} \sim^{iid} N(0, 0.1574)$ and $\varepsilon_{t+n}^{*} = \sum_{i=1}^{n} \phi^{n-i}\theta_{t+i}$ where $\theta_t \sim^{iid} N(0,1)$ and n is the length of the futures contract.

Table 5: Bubble Date Stamping (weekly frequency)

Series	*BSADF*	$\mathbf{BSADF}^*_{wb}$	**Series**	*BSADF*	$\mathbf{BSADF}^*_{wb}$
$s_{t+1} - f_{t,1}$	-	1997M4-2019M7	$s_{t+15} - f_{t,15}$	2008M4-2008M5	2007M12-2008M9
	-			2014M12-2015M3	2014M11-2016M3
$s_{t+3} - f_{t,3}$	-	1997M6-2019M7	$s_{t+18} - f_{t,18}$	2007M1-2007M2	2006M9-2007M11
	-			2008M5-2008M7	2008M3-2008M8
				2014M12-2015M4	2014M11-2015M4
$s_{t+6} - f_{t,6}$	2008M10-2009M3	2007M6-2019M7	$s_{t+21} - f_{t,21}$	2007M6-2007M10	2007M6-2008M2
				-	2015M11-2015M12
$s_{t+9} - f_{t,9}$	-	2006M9-2008M10	$s_{t+24} - f_{t,24}$	-	2007M9-2007M10
	2008M11-2009M3	2008M10-2009M9		-	2015M9-2015M10
	2014M11-2015M4	2014M10-2015M9			
$s_{t+12} - f_{t,12}$	-	2007M4-2007M9			
	2008M12-2009M1	2008M12-2011M11			
	2014M12-2015M4	2014M10-2016M3			

Note: Bubble date stamping application on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. We compute right-tailed finite sample critical values for both tests using $10,000$ Monte Carlo for the BSADF test and $9,999$ bootstrap replications for the BSADF^*_{wb} test. Both tests are performed at a 5% level of significance and a constant is included in the regression.

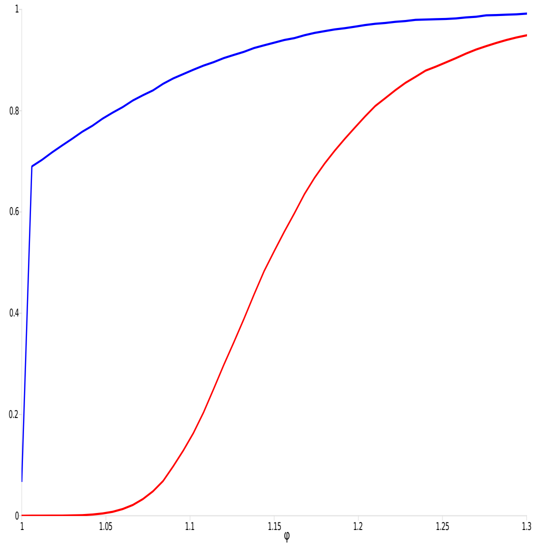
Table 6: Bubble Date Stamping (monthly frequency)

Series	*BSADF*	$\mathbf{BSADF}^{*}_{wb}$	**Series**	*BSADF*	$\mathbf{BSADF}^{*}_{wb}$
$s_{t+1} - f_{t,1}$	-	1998M6-2009M6	$s_{t+15} - f_{t,15}$	-	2007M7-2007M11
	-	2009M9-2019M6		-	2015M11-2015M12
$s_{t+3} - f_{t,3}$	-	2007M3-2009M1	$s_{t+18} - f_{t,18}$	-	2007M7-207M11
	-	2015M1-2015M7		-	2015M9-2015M12
$s_{t+6} - f_{t,6}$	-	2009M4-2009M6	$s_{t+21} - f_{t,21}$	2007M6-2007M10	2007M6-2008M2
	2015M6-2015M7	2015M4-2015M8		-	2015M11-2015M12
$s_{t+9} - f_{t,9}$	-	2006M4-2008M4	$s_{t+24} - f_{t,24}$	-	2007M9-2007M10
	-	2015M9-2015M12		-	2015M9-2015M10
$s_{t+12} - f_{t,12}$	-	2007M6-2007M11			
		2015M9-2015M10			

Note: Bubble date stamping application on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a monthly frequency, constituting 287 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. We compute right-tailed finite sample critical values for both tests using $10,000$ Monte Carlo for the BSADF test and $9,999$ bootstrap replications for the BSADF^{*}_{wb} test. Both tests are performed at a 5% level of significance and a constant is included in the regression.

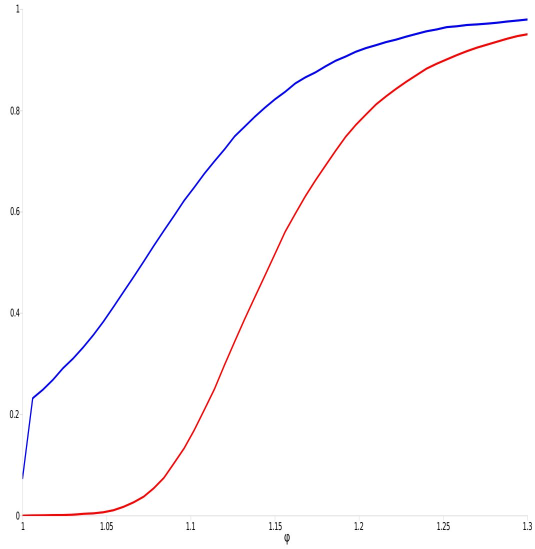
3.10 Figures

Figure 3: Finite Sample Power



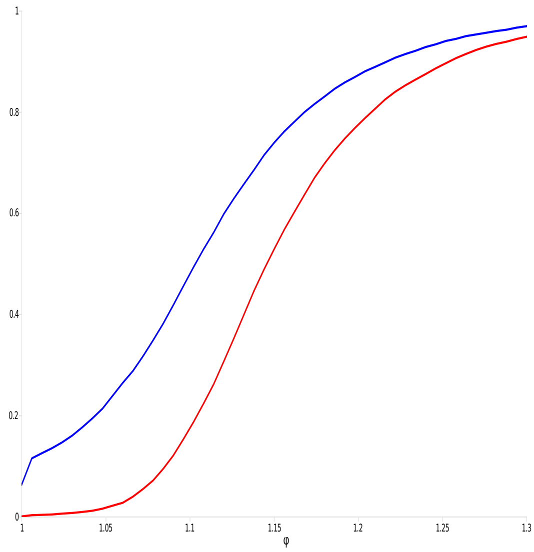
(a)

1 Month Futures Contract



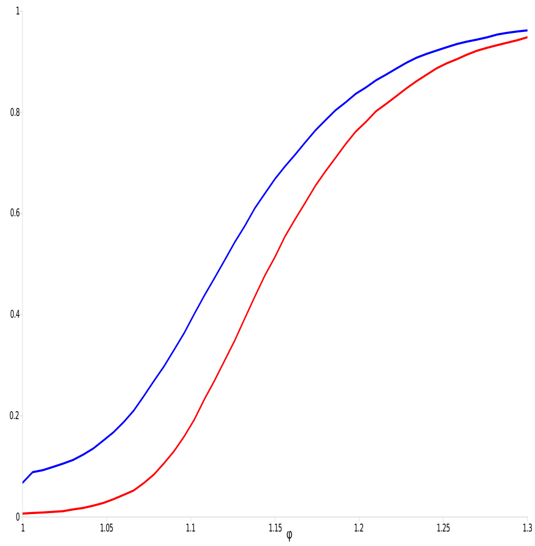
(b)

3 Month Futures Contract



(c)

6 Month Futures Contract



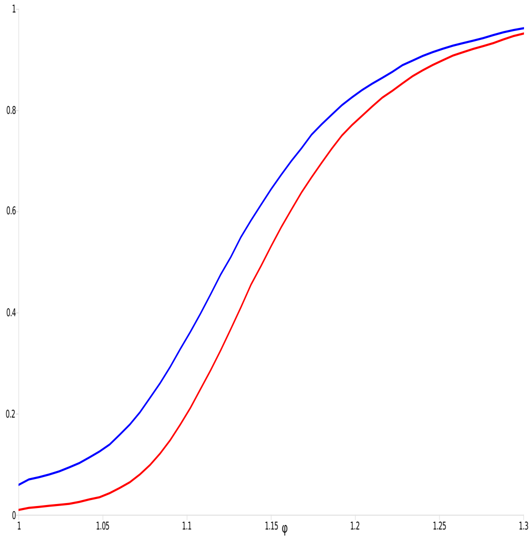
(d)

9 Month Futures Contract

BSADF test: ——— BSADF$^*_{wb}$ test: ———

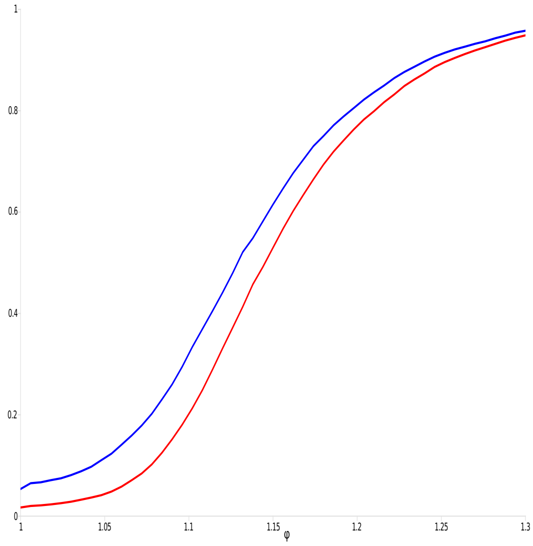
Data generated according to $y_{t+n} = \varepsilon^*_{t+n}$, $t = 1, ..., 180$ and $y_{t+n} = b_t + \varepsilon_{t+n} + \varepsilon^*_{t+n}$, $t = 181, ..., 200$ where $b_t = \phi_t b_{t-1} + \varepsilon_t$, $t = 1, ..., 200$, $\varepsilon^*_t \sim^{iid} N(0, 0.1574)$ and $\varepsilon_t \sim^{iid} N(0, 0.3^2)$.

Figure 4: Finite Sample Power



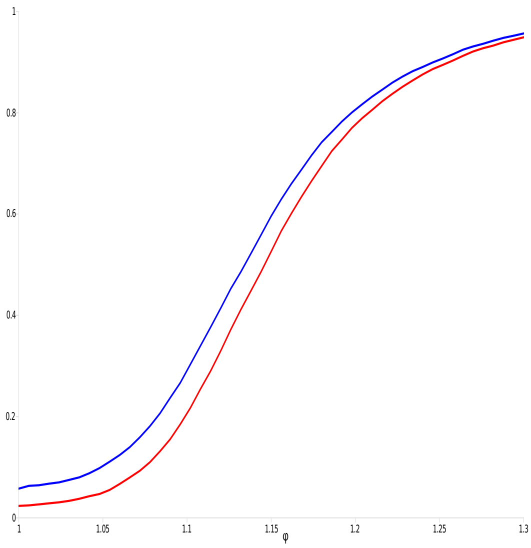
(a)

12 Month Futures Contract



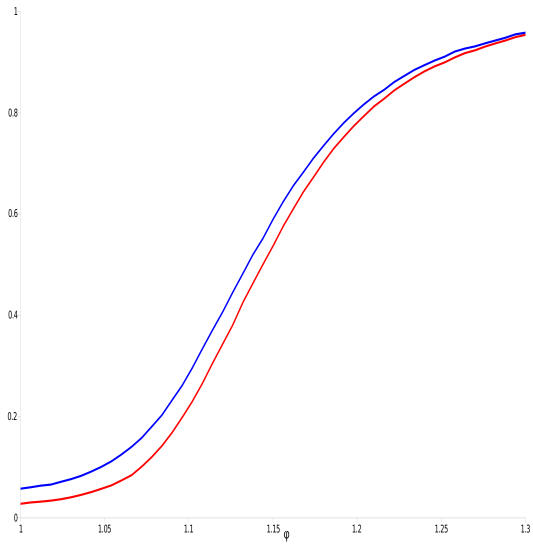
(b)

15 Month Futures Contract



(c)

18 Month Futures Contract



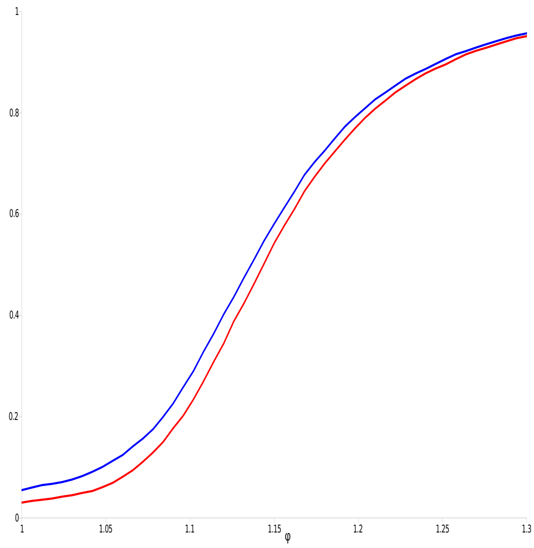
(d)

21 Month Futures Contract

BSADF test: ——— BSADF$^*_{wb}$ test: ———

Data generated according to $y_{t+n} = \varepsilon^*_{t+n}$, $t = 1, ..., 180$ and $y_{t+n} = b_t + \varepsilon_{t+n} + \varepsilon^*_{t+n}$, $t = 181, ..., 200$ where $b_t = \phi_t b_{t-1} + \varepsilon_t$, $t = 1, ..., 200$, $\varepsilon^*_t \sim^{iid} N(0, 0.1574)$ and $\varepsilon_t \sim^{iid} N(0, 0.3^2)$.

Figure 5: Finite Sample Power



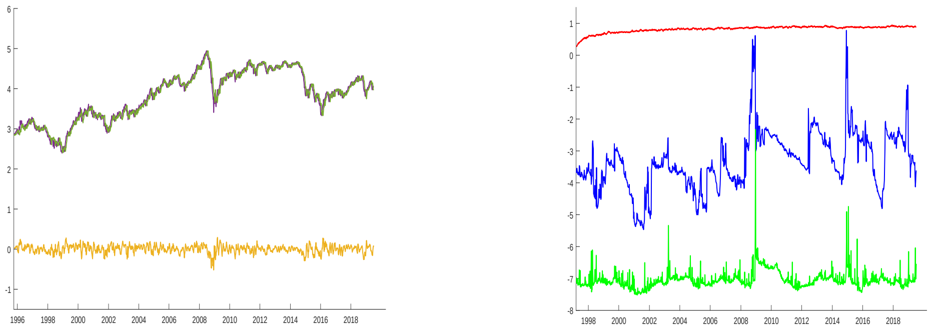
(a)

24 Month Futures Contract

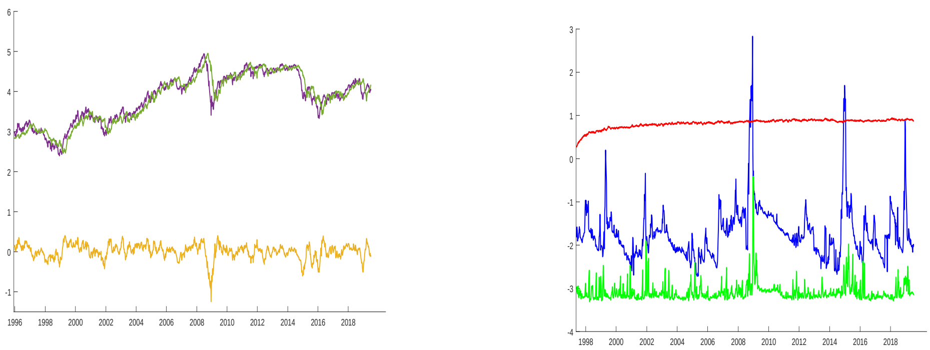
BSADF test: ——— $\quad$ BSADF$^*_{wb}$ test: ———

Data generated according to $y_{t+n} = \varepsilon^*_{t+n}$, $t = 1, ..., 180$ and $y_{t+n} = b_t + \varepsilon_{t+n} + \varepsilon^*_{t+n}$, $t = 181, ..., 200$ where $b_t = \phi_t b_{t-1} + \varepsilon_t$, $t = 1, ..., 200$, $\varepsilon^*_t \sim^{iid} N(0, 0.1574)$ and $\varepsilon_t \sim^{iid} N(0, 0.3^2)$.

Figure 6: Bubble Date Stamping (weekly frequency)



(a) 1 Month Futures Contract $(s_{t+1} - f_{t,1})$



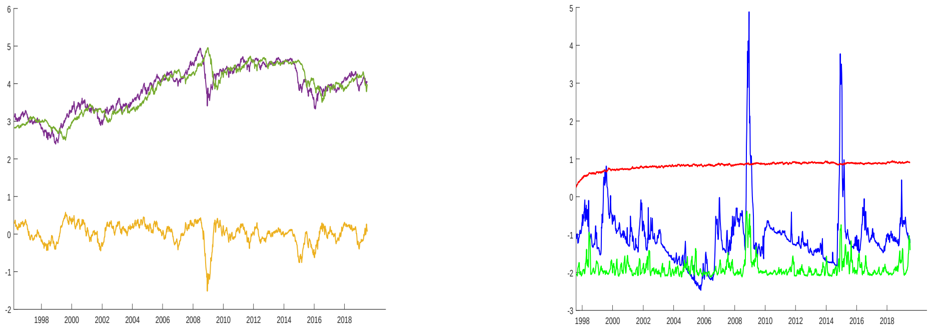
(b) 3 Month Futures Contract $(s_{t+3} - f_{t,3})$

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

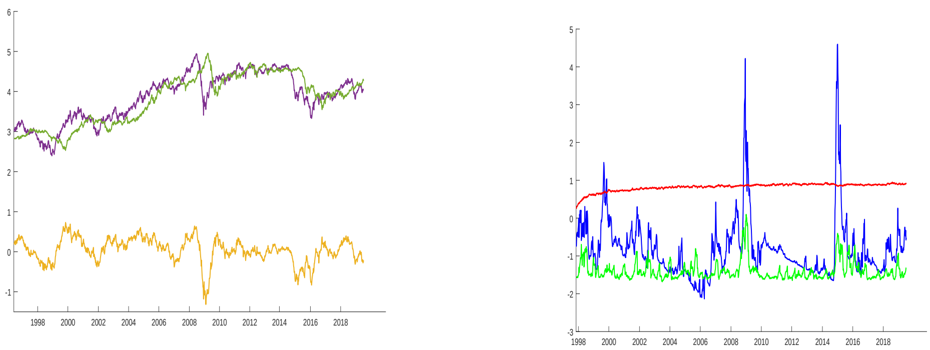
BSADF test-statistics: —— BSADF$^{*}_{wb}$ critical values: ——
BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them $(s_{t+n} - f_{t,n})$, where n is the time to maturity at weekly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using $10,000$ Monte Carlo for the BSADF test and $9,999$ bootstrap replications for the BSADF$^{*}_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 7: Bubble Date Stamping (weekly frequency)



(a) 6 Month Futures Contract $(s_{t+6} - f_{t,6})$



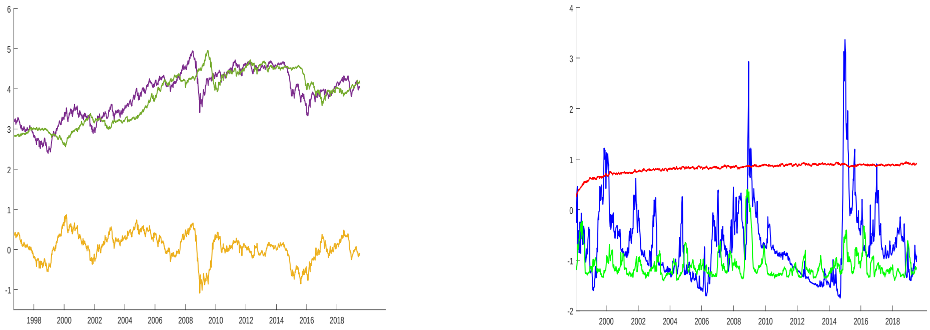
(b) 9 Month Futures Contract $(s_{t+9} - f_{t,9})$

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

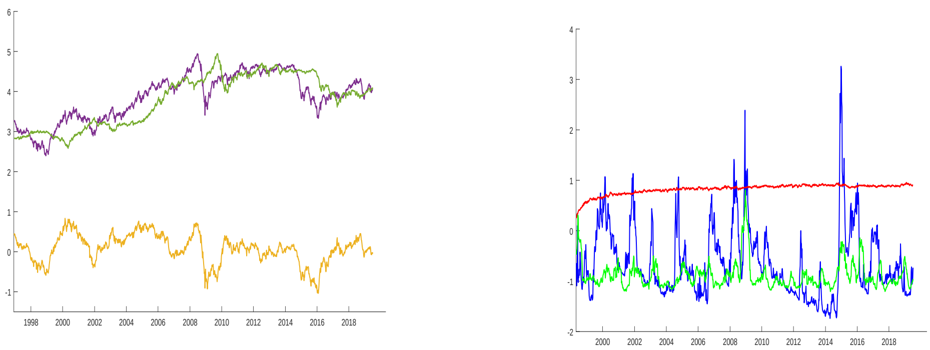
BSADF test-statistics: —— BSADF$^{*}_{wb}$ critical values: —— BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them $(s_{t+n} - f_{t,n})$, where n is the time to maturity at weekly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using 10,000 Monte Carlo for the BSADF test and 9,999 bootstrap replications for the BSADF$^{*}_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 8: Bubble Date Stamping (weekly frequency)



(a) 12 Month Futures Contract ($s_{t+12} - f_{t,12}$)



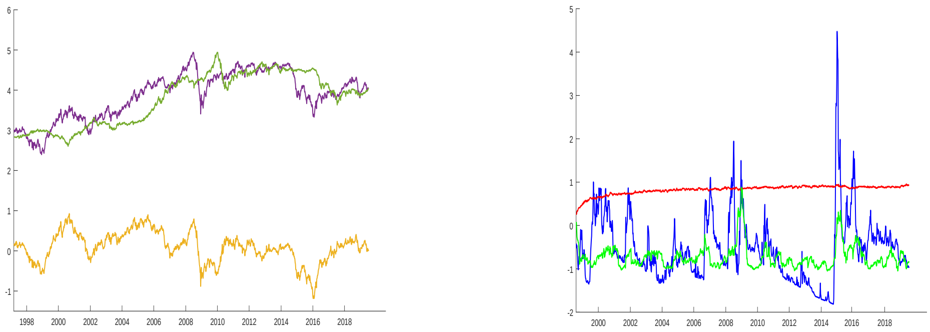
(b) 15 Month Futures Contract ($s_{t+15} - f_{t,15}$)

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

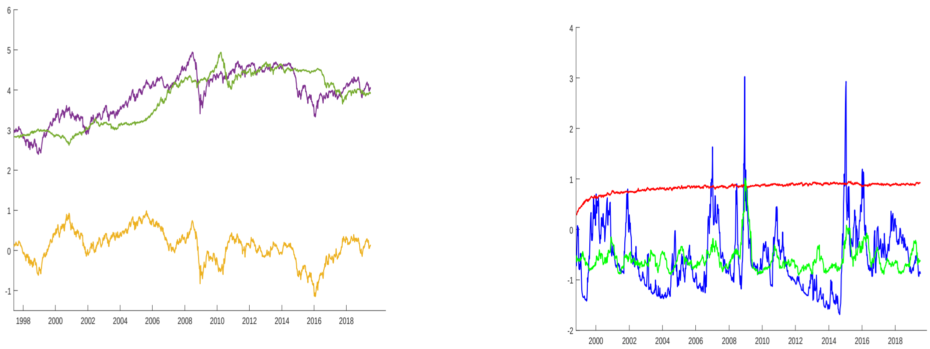
BSADF test-statistics: —— BSADF$^*_{wb}$ critical values: —— BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them ($s_{t+n} - f_{t,n}$), where n is the time to maturity at weekly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using 10,000 Monte Carlo for the BSADF test and 9,999 bootstrap replications for the BSADF$^*_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 9: Bubble Date Stamping (weekly frequency)



(a) 18 Month Futures Contract ($s_{t+18} - f_{t,18}$)



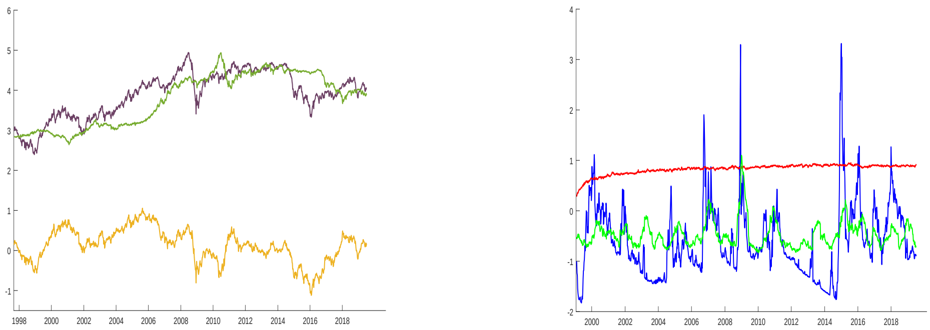
(b) 21 Month Futures Contract ($s_{t+21} - f_{t,21}$)

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

BSADF test-statistics: —— BSADF$^*_{wb}$ critical values: ——
BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them ($s_{t+n} - f_{t,n}$), where n is the time to maturity at weekly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using 10,000 Monte Carlo for the BSADF test and 9,999 bootstrap replications for the BSADF$^*_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 10: Bubble Date Stamping (weekly frequency)



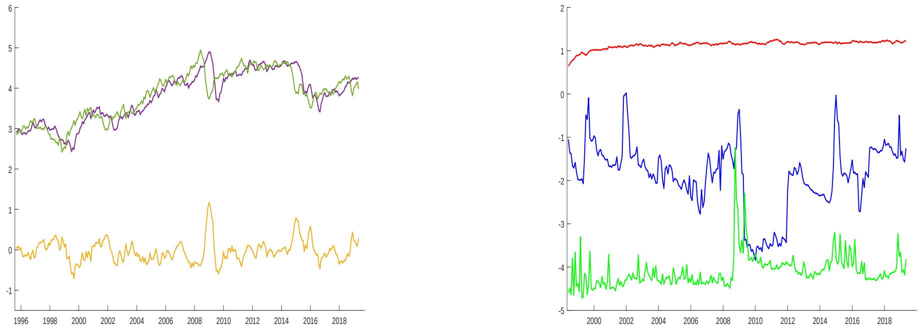
(a) 24 Month Futures Contract ($s_{t+24} - f_{t,24}$)

spot price series: ——— futures contract price series: ——— spot price - futures contract price series: ———

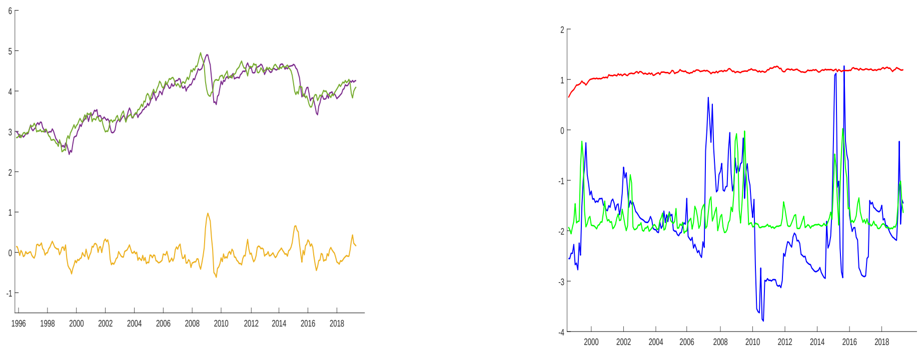
BSADF test-statistics: ——— BSADF$^{*}_{wb}$ critical values: ———
BSADF critical values: ———

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them ($s_{t+n} - f_{t,n}$), where n is the time to maturity at weekly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using $10,000$ Monte Carlo for the BSADF test and $9,999$ bootstrap replications for the BSADF$^{*}_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 11: Bubble Date Stamping (monthly frequency)



(a) 1 Month Futures Contract $(s_{t+1} - f_{t,1})$



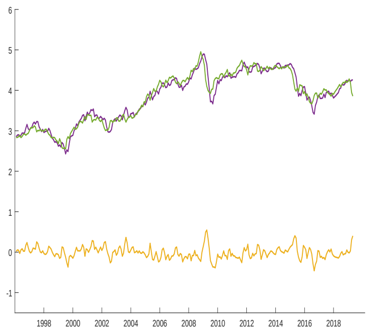
(b) 3 Month Futures Contract $(s_{t+3} - f_{t,3})$

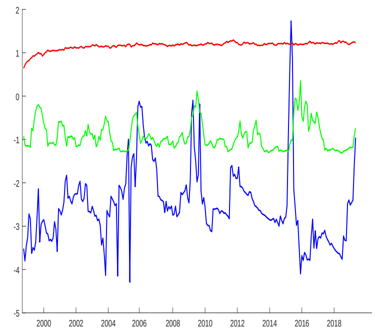
spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

BSADF test-statistics: —— BSADF$^*_{wb}$ critical values: —— BSADF critical values: ——

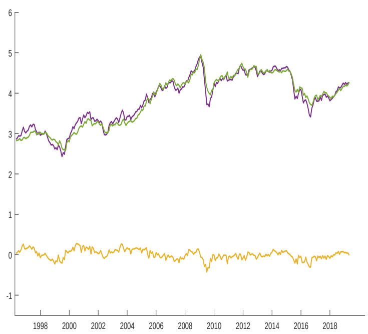
Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them $(s_{t+n} - f_{t,n})$, where n is the time to maturity at monthly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using 10,000 Monte Carlo for the BSADF test and 9,999 bootstrap replications for the BSADF$^*_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

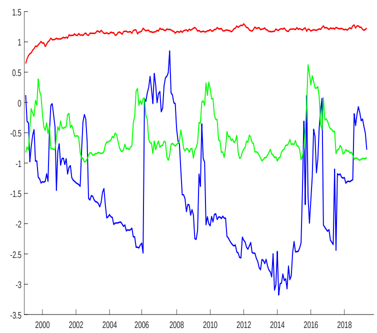
Figure 12: Bubble Date Stamping (monthly frequency)





(a) 6 Month Futures Contract $(s_{t+6} - f_{t,6})$





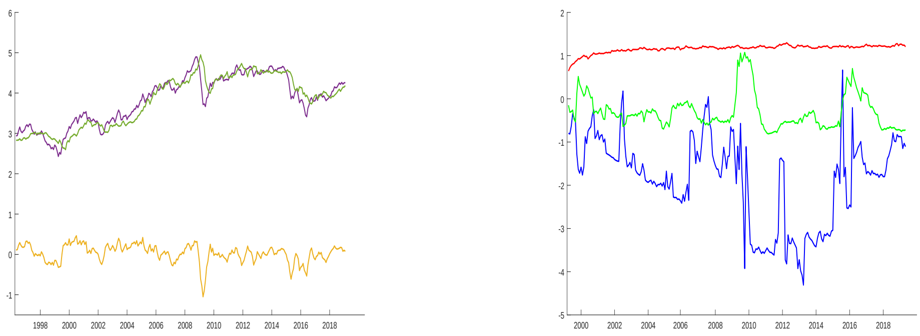
(b) 9 Month Futures Contract $(s_{t+9} - f_{t,9})$

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

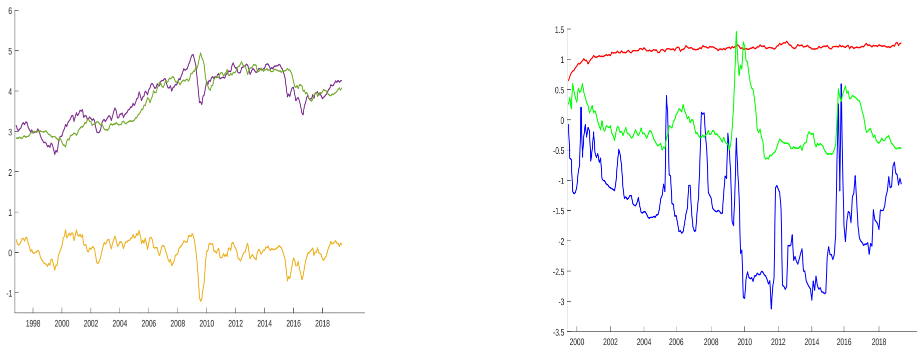
BSADF test-statistics: —— BSADF$^{*}_{wb}$ critical values: ——
BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them $(s_{t+n} - f_{t,n})$, where n is the time to maturity at monthly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using $10,000$ Monte Carlo for the BSADF test and $9,999$ bootstrap replications for the BSADF$^{*}_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 13: Bubble Date Stamping (monthly frequency)



(a) 12 Month Futures Contract ($s_{t+12} - f_{t,12}$)



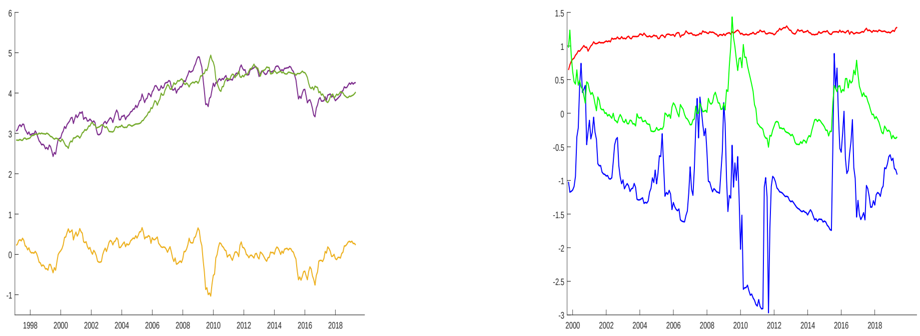
(b) 15 Month Futures Contract ($s_{t+15} - f_{t,15}$)

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

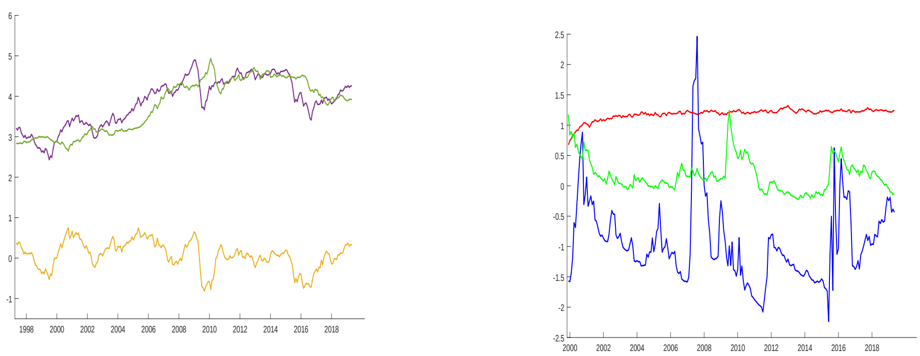
BSADF test-statistics: —— BSADF$^{*}_{wb}$ critical values: —— BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them ($s_{t+n} - f_{t,n}$), where n is the time to maturity at monthly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using 10,000 Monte Carlo for the BSADF test and 9,999 bootstrap replications for the BSADF$^{*}_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 14: Bubble Date Stamping (monthly frequency)



(a) 18 Month Futures Contract $(s_{t+18} - f_{t,18})$



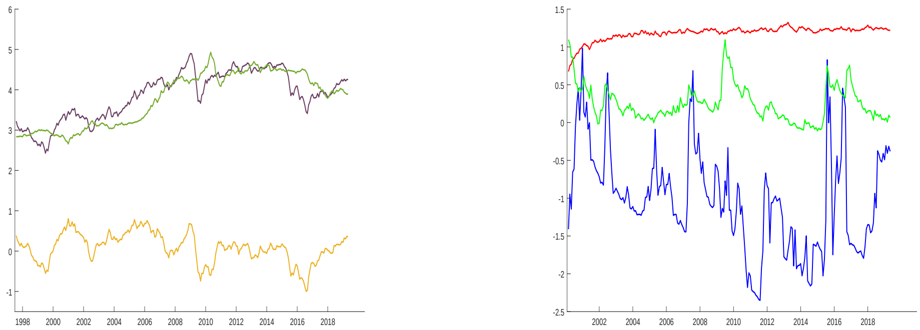
(b) 21 Month Futures Contract $(s_{t+21} - f_{t,21})$

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

BSADF test-statistics: —— BSADF$^*_{wb}$ critical values: —— BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them $(s_{t+n} - f_{t,n})$ where n is the time to maturity at monthly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using 10,000 Monte Carlo for the BSADF test and 9,999 bootstrap replications for the BSADF$^*_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

Figure 15: Bubble Date Stamping (monthly frequency)



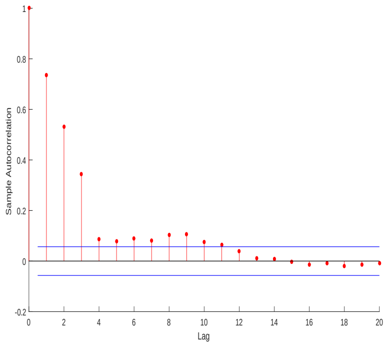
(a) 24 Month Futures Contract ($s_{t+24} - f_{t,24}$)

spot price series: —— futures contract price series: —— spot price - futures contract price series: ——

BSADF test-statistics: —— BSADF$^*_{wb}$ critical values: —— BSADF critical values: ——

Notes:
Left-hand side: WTI crude oil future spot and futures contract logarithmic prices and the difference between them ($s_{t+n} - f_{t,n}$), where n is the time to maturity at monthly frequency from September 1995 to July 2019.
Right-hand side: Right-tail finite sample critical values are simulated for both tests using $10,000$ Monte Carlo for the BSADF test and $9,999$ bootstrap replications for the BSADF$^*_{wb}$ test. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Both tests are performed at a 5% level of significance and a constant is included in the regressions.

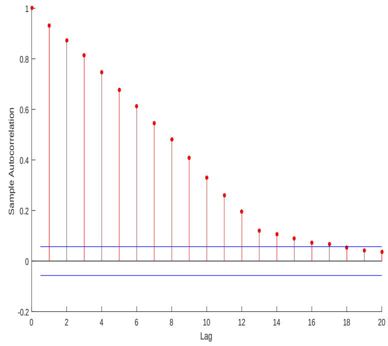
Figure 16: Sample Autocorrelation Function (weekly frequency)



(a)

1 Month Futures Contract

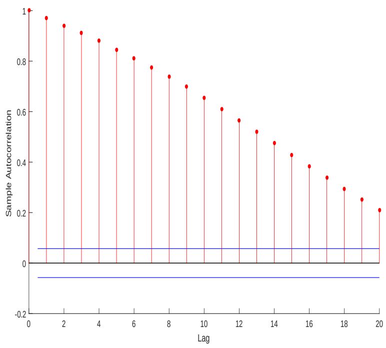
$s_{t+1} - f_{t,1}$



(b)

3 Month Futures Contract

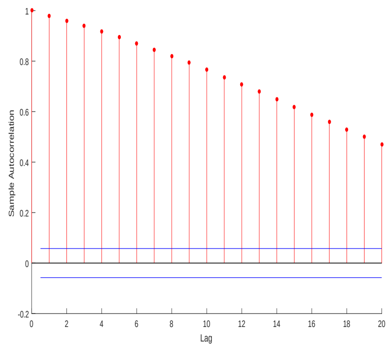
$s_{t+3} - f_{t,3}$



(c)

6 Month Futures Contract

$s_{t+6} - f_{t,6}$



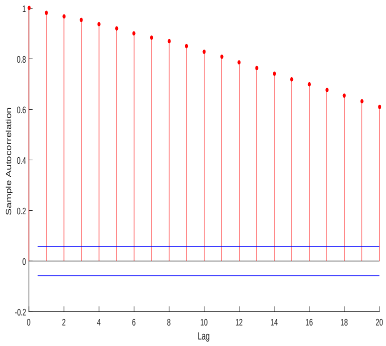
(d)

9 Month Futures Contract

$s_{t+9} - f_{t,9}$

Note: Sample autocorrelation functions on the logarithmic differences between the WTI crude oil future spot and futures contract prices, $s_{t+n} - f_{t,n}$, where n is the time to maturity for the period September 1995 to July 2019 at weekly frequency.

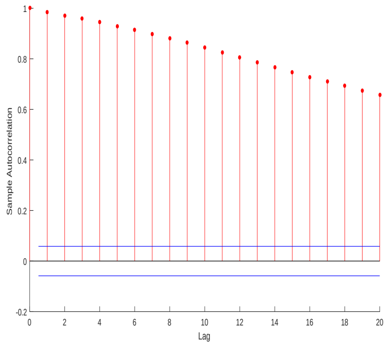
Figure 17: Sample Autocorrelation Function (weekly frequency)



(a)

12 Month Futures Contract

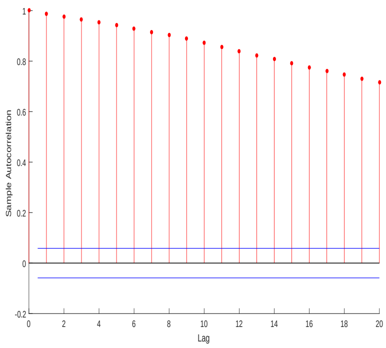
$s_{t+12} - f_{t,12}$



(b)

15 Month Futures Contract

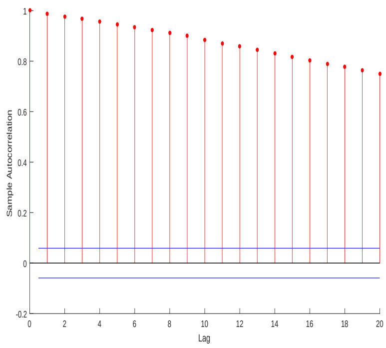
$s_{t+15} - f_{t,15}$



(c)

18 Month Futures Contract

$s_{t+18} - f_{t,18}$



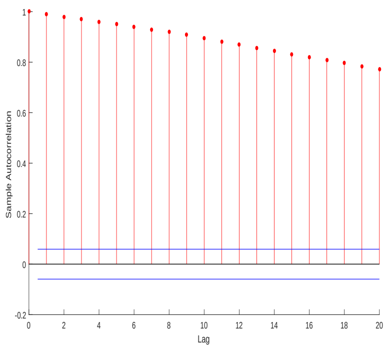
(d)

21 Month Futures Contract

$s_{t+21} - f_{t,21}$

Note: Sample autocorrelation functions on the logarithmic differences between the WTI crude oil future spot and futures contract prices, $s_{t+n} - f_{t,n}$, where n is the time to maturity for the period September 1995 to July 2019 at weekly frequency.

Figure 18: Sample Autocorrelation Function (weekly frequency)



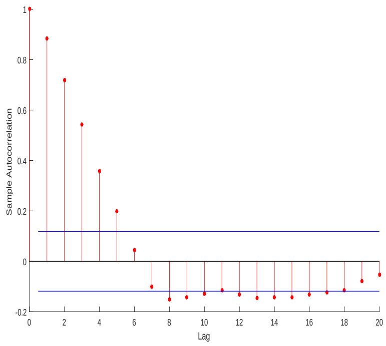
(a)

24 Month Futures Contract

$$s_{t+24} - f_{t,24}$$

Note: Sample autocorrelation functions on the logarithmic differences between the WTI crude oil future spot and futures contract prices, $s_{t+n} - f_{t,n}$, where n is the time to maturity for the period September 1995 to July 2019 at weekly frequency.

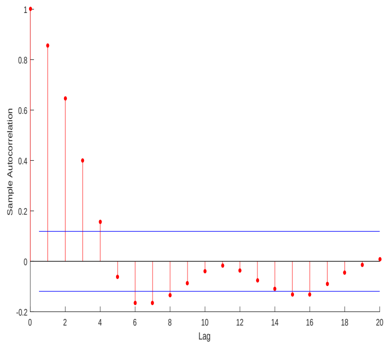
Figure 19: Sample Autocorrelation Function (monthly frequency)



(a)

1 Month Futures Contract

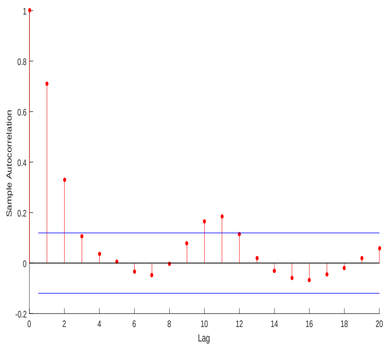
$s_{t+1} - f_{t,1}$



(b)

3 Month Futures Contract

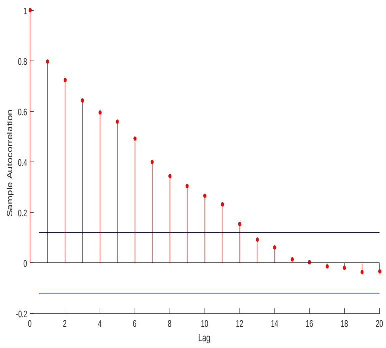
$s_{t+3} - f_{t,3}$



(c)

6 Month Futures Contract

$s_{t+6} - f_{t,6}$



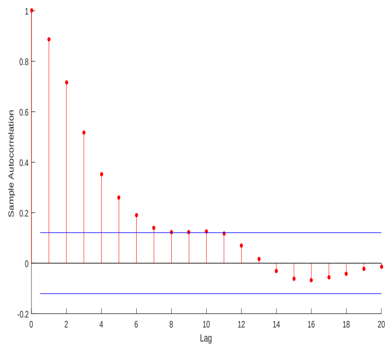
(d)

9 Month Futures Contract

$s_{t+9} - f_{t,9}$

Note: Sample autocorrelation functions on the logarithmic differences between the WTI crude oil future spot and futures contract prices, $s_{t+n} - f_{t,n}$, where n is the time to maturity for the period September 1995 to July 2019 at monthly frequency.

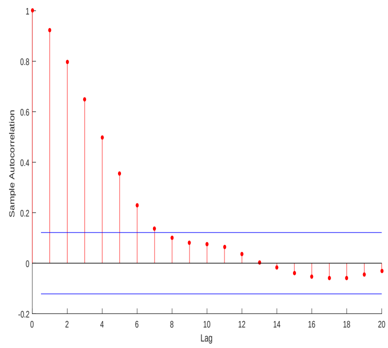
Figure 20: Sample Autocorrelation Function (monthly frequency)



(a)

12 Month Futures Contract

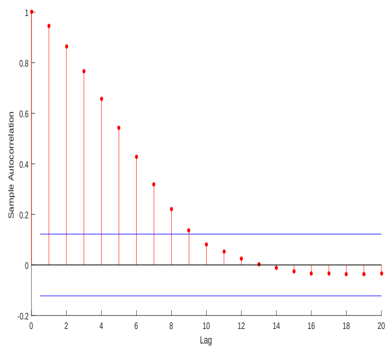
$s_{t+12} - f_{t,12}$



(b)

15 Month Futures Contract

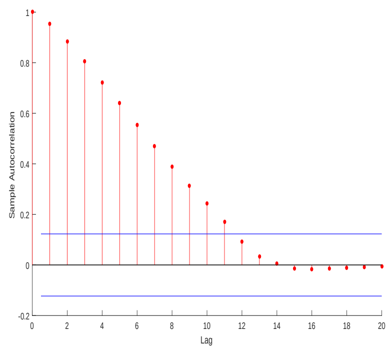
$s_{t+15} - f_{t,15}$



(c)

18 Month Futures Contract

$s_{t+18} - f_{t,18}$



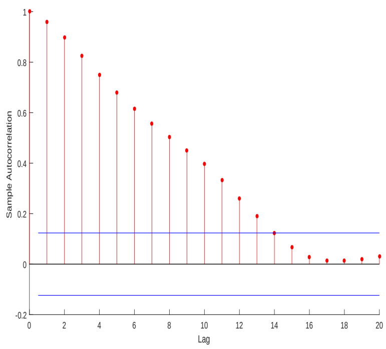
(d)

21 Month Futures Contract

$s_{t+21} - f_{t,21}$

Note: Sample autocorrelation functions on the logarithmic differences between the WTI crude oil future spot and futures contract prices, $s_{t+n} - f_{t,n}$, where n is the time to maturity for the period September 1995 to July 2019 at monthly frequency.

Figure 21: Sample Autocorrelation Function (monthly frequency)



(a)

24 Month Futures Contract

$$s_{t+24} - f_{t,24}$$

Note: Sample autocorrelation functions on the logarithmic differences between the WTI crude oil future spot and futures contract prices, $s_{t+n} - f_{t,n}$, where n is the time to maturity for the period September 1995 to July 2019 at monthly frequency.

4 Testing for Bubbles in Commodity Spot and Futures Using a Co-explosive Autoregression

4.1 Introduction

It has been widely acknowledged that standard econometric analysis does not allow for explosive behaviour, making bubble identification rather challenging. Econometric tests on bubble identification such as unit root and cointegration tests are well documented in the literature. In this chapter we study asset bubbles under the assumption that rational expectations hold, however the tests and their extensions can be equally useful not only in identifying rational bubbles but intrinsic bubbles -bubbles that depend on market fundamentals (Froot and Obstfeld 1992)- or explosive bubbles -bubbles with high probability of bursting- as well, see *inter alia* Evans(1991).[16]

4.1.1 Cointegration in a Rational Expectations Framework

In a rational bubble regime, asset prices move away from their market fundamentals and therefore the equilibrium condition is violated. Non-stationary (or non-mean-reverting) deviations from equilibrium may signal the indication of a bubble in the long-run. Cointegration analysis is considered as one of the main approaches in testing for deviations from equilibrium. For instance, Campbell and Shiller (1987) introduce the argument that cointegration between stock prices and dividends can be considered as evidence of no bubble. In their empirical work, Campbell and Shiller (1987) argue that there is persistence on the deviations from fundamentals, however quite sensitive to the discount rate. In an extension of Campbell and Shiller's (1987) cointegrating restriction, Craine (1993) imposes a robust no-rational bubble constraint that does not hold the assumption of a constant discount rate or a particular asset pricing model and provides evidence on the non-stationarity of the discount factor for the S&P 500, arguing that any statistical inference on bubble existence might be rather inconclusive.

An alternative hypothesis of a rational bubble is firstly introduced by West (1987) who examines the effect of the market fundamental on the asset price and concludes that the impact of dividends on asset prices can be either attributed to asset pricing

[16]Rational expectations bubbles refer to the scenario where traders with rational expectations extrapolate current blips in the asset markets into forming expectations about higher asset prices in the future.

model misspecification or bubble. The above argument is subject to criticism as rejections of the alternative hypothesis might be justified by other reasons such as the inadequacy of the model to explain the bubble episode (Flood et al. 1994). Arguably, conventional univariate econometric tests can provide misleading inference due to omitted variable biases (Flood and Garber 1980), model misspecifications or inconsistent statistical tests (Flood and Hodrick 1986), low power on identifying periodically collapsing bubble episodes (Evans 1991), and size distortions and low power (see *inter alia* Stock 1991 and Campbell and Perron 1991) leading to false rejection of the null hypothesis of no bubble.

In their seminal paper Diba and Grossman (1988) emphasize the importance of unit root testing in rational bubble identification by introducing a left-tailed unit root process to test the null hypothesis of a unit root assuming a time-invariant discount rate. Furthermore, they argue that if both stock prices and dividends are stationary in differences then there is no evidence of rational bubble. In their empirical exercise, they test the real S&P 500 stock price index between 1871 and 1986 to conclude that stock prices and dividends are stationary in differences and therefore there is no evidence of rational bubbles on the S&P 500. In a cointegration framework, Diba and Grossman (1988) consider the Bharghava (1986) ratios to infer that if two series e.g. the stock price and the dividend are cointegrated then there is no evidence of bubble, which is criticised by Evans (1991) on the basis of the complexity of the bubble characteristics and the non-linearity of bubbles that cannot be captured by conventional cointegration tests.

In a cointegrated vector autoregression framework, Johansen and Swensen (1999), Johansen and Swensen (2004) and Johansen and Swensen (2011) test the restrictions of rational expectations as firstly proposed by Baillie (1989), although following a different methodology to account for non-stationarity. Furthermore, they generate the likelihood ratio tests under the restrictions of the rational expectation hypothesis and argue that present value models pose restrictions on the cointegrating relationships.

4.1.2 Analysis of Coexplosive and Cointegrated Processes

As an extension of the models suggested by Campbell and Shiller (1987) and Campbell and Shiller (1988) and Johansen and Swensen (1999), Johansen and Swensen (2004) and

Johansen and Swensen (2011), Nielsen (2010) introduces the idea of co-explosiveness to allow the standard cointegrated VAR models to test for the existence of bubbles. In particular, Nielsen (2010) proposes a VAR model that allows both unit roots and explosive characteristic roots, utilising the standard cointegration techniques introduced in Johansen (1991). The coexplosive and cointegrated vector autoregressive model arises as a restriction to the standard VAR model and allows both a random walk and an explosive stochastic component with a characteristic root, $\rho > 1$.

This model contradicts Diba and Grossman (1988) on the fact that two series can be cointegrated and yet, their linear combination contain an explosive component (Engsted 2006). As a result, the VAR approach developed by Johansen (1991) offers the advantage of testing for cointegration while simultaneously testing whether at least one of the variables has an explosive characteristic root since testing for the number of cointegrating vectors in the coexplosive case is similar to the standard Johansen procedure. The reason for this is that the asymptotic distribution of the likelihood ratio test when there is an explosive root is the same as in the standard Johansen cointegration test (Nielsen 2010). In his empirical exercise, Nielsen (2010) applies the likelihood ratio test in Yugoslavian hyperinflation data of 1990-1994 and concludes that prices and money supply contain both an explosive and random walk component whereas their linear combination is a unit root process.

To test the rational bubble hypothesis, Engsted and Nielsen (2012) apply the likelihood ratio test on US real stock price and dividend data between 1974 and 2000 to conclude that the real stock prices contain an explosive characteristic root, however the evidence is rather weak as the null hypothesis of no cointegrating relationship is marginally rejected. In the same framework, Engsted (2006) extends the Diba and Grossman (1988) dataset to cover the period 1871-2000 (whereas the original dataset only covers the period between 1871 and 1986) and applies the Johansen (1991) methodology to test for the long-term relationship between stock prices and dividends. Engsted (2006) comes to a conclusion that is consistent with the standard present value model with explosive bubbles as described in Diba and Grossman (1988a) and Diba and Grossman (1988) and infers that stock prices contain both a unit root and an explosive root with dividends being a unit root process and their linear relationship containing an explosive root.

4.1.3 Recent Unit Root Tests for Explosive Behaviour

In standard time-series econometrics, non-stationary variables are either treated as first-order integrated, second-order integrated or fractionally integrated. However, fractional integration techniques have been applied to bubble testing as well. Cuñado et al. (2005) suggest a fractionally integrated approach in NASDAQ stock prices and dividends and provide evidence that bubble detection can be rather sensitive to the sampling frequency of the data. Particularly, they find that daily and weekly frequency data suggest fractional cointegration whereas monthly frequency data suggest no cointegration. Possible explanations for this inconclusive inference can be the bias due to the low frequency (temporal aggregation problem) or the sample size (Cuñado et al. 2005).

In an empirical study, Koustas and Serletis (2005) apply fractional integration methodologies to the logarithmic dividend yield of the S&P 500 finding evidence in support of a rational bubble. However, Koustas and Serletis (2005) technique is subject to criticism by Frömmel and Kruse (2012) as it does not account for structural breaks. In contrast, Frömmel and Kruse (2012) suggest a different fractional integration approach that considers structural breaks and changes in persistence in line with Sibbertsen and Kruse (2009).

The problem of asymptotic bias of integrated, near-integrated or explosive regressors has been emphasised by Magdalinos and Phillips (2009) who extent Phillips and Magdalinos (2008) asymptotic theory for fully explosive cointegrated regressors to account for moderately explosive regressors. Furthermore, they suggest that the relationship between the explosive regressors determines the limit behaviour of the least squares estimator and that in a moderately explosive framework the regressors result on a mixed normal limit.

More recent research focuses on right-tailed unit root processes, testing the alternative hypothesis of explosive behaviour. Phillips et al. (2011) argue that explosive behaviour in the asset prices but not in the market fundamentals might be perceived as a bubble episode. Moreover, Phillips et al. (2011) propose a forward recursive right-tailed supremum Augmented Dickey-Fuller (SADF) test useful on identifying bubbles as they grow and they apply the SADF test on NASDAQ stock prices and dividends to find evidence of the dot-com bubble. The test has been successfully applied to commodity future prices (Gilbert 2010), commodity and house prices (Homm and Breitung 2012) and the exchange rate market (Bettendorf and Chen 2013).

The Phillips et al. (2011) methodology seems to be less powerful in a multiple bubble regime, indicating pseudo-stationarity issues. For this reason, Phillips et al. (2015) suggest a backward and a generalised version of the SADF test capable of identifying historical and real time multiple bubble episodes with periodically collapsing behaviour, the BSADF and the GSADF test respectively. The GSADF test of Phillips et al. (2015) appears to have good size and power properties since it has been designed to run recursively for different starting and ending points of the sample in a way that can detect multiple explosive episodes. The BSADF test is widely used for date-stamping past bubble episodes since it can precisely estimate the origination and termination dates of the bubbles, being quite powerful on identifying bubbles that occur at the end of the sample. As an empirical application, Phillips et al. (2015) apply the proposed GSADF test to the S&P 500 stock price and dividend index for the period January 1871 - December 2010 and find strong evidence of bubble behaviour. Then, to date-stamp the explosive episodes, they apply the BSADF test that successfully manages to identify more than six historical banking crises and bubble episodes from January 1871 to December 2010.

Tsvetanov et al. (2016) apply the Phillips et al. (2015) methodology to test for bubble episodes to crude oil spot and future markets. Specifically, Tsvetanov et al. (2016) use crude oil prices for spot and futures contracts on NYMEX from 1995 to 2013 to test for explosive behaviour to find evidence of explosiveness at weekly and monthly frequency between 2004 and 2008. The most important finding of their study is that longer-dated contracts suggest that the origination date of the bubble episode is earlier compared to shorter-dated contracts or even spot prices and they provide strong evidence against the null hypothesis of no bubble in support of their conclusion that the evidence of bubble existence becomes stronger as maturity increases. In particular, twelve, fifteen, eighteen, twenty one and twenty four month futures contracts suggest that there is a multiple or continuous bubble episode that starts in early 2004 concluding that futures contracts with maturity over six months have been significantly overpriced above their fundamentals, since the beginning of 2004. Their empirical findings seem to be in accordance to the related literature as increased investment flows into commodity derivatives markets inflated the oil futures contract prices (Sockin and Xiong 2015, Tang and Xiong 2012 and Singleton 2014).

Extending Phillips et al. (2015) methodology, Phillips and Shi (2018) account for a delay bias that seems to affect the date-stamping outcome. To do so, they introduce a reverse sample regression into the recursive window methodology of Phillips et al. (2015) and they incorporate a market recovery parameter to date when asset prices return to equilibrium. Furthermore, they distinguish the crashes of bubble episodes into sudden, disturbing and smooth, differentiating the crash date from the market correction date. Applying their model on the NASDAQ stock market index between January 1973 and August 2013, Phillips and Shi (2017) identify four stages of the dot-com bubble episode; the origination date, the implosion date, the market correction or recovery date and finally a further correction date resulting in a more precise real-time mechanism for bubble identification.

More recently, Pavlidis et al. (2017) argue that periodically collapsing bubbles result in a disruption of the relationship between future spot and futures contract prices and therefore market efficiency. As a consequence, futures contract prices become a biased predictor of the future spot prices with the degree of bias being explosive. In their empirical application of the Phillips et al. (2015) tests and the rolling Fama regressions to the "German Hyperinflation" period (December 1921 to August 1923), the "Recent Float" period (January 1979 to December 2013) and the U.S equity market (December 1982 to March 2015), Pavlidis et al. (2017) conclude that explosive behaviour in asset prices does not necessarily imply the existence of a bubble episode as explosiveness might be attributed to bubble behaviour in fundamentals. Thus, any statistical inference on rational bubbles based on unit root testing can be rather inconclusive as any rejection of the null hypothesis of a unit root can be either attributed to the existence of bubbles or misspecification of the asset pricing model or both.

In the same framework, Pavlidis et al. (2018) apply their proposed methodology to the crude oil market and use market expectations instead of futures contract prices, as the former is not influenced by the risk premium, to test for speculative bubbles on WTI crude oil during the period January 1990 to December 2013. They support that since the financialisation of the oil futures markets in 2003, it is the fundamentals that drove oil prices up and not the the development of speculative bubble and that explosive episodes in the oil market should be perceived as changes in fundamentals rather than evidence of speculative bubbles.

In this article we consider the application of Nielsen (2010) approach to test for cointegrating relationships across different series while simultaneously testing whether the series contain any explosive components, allowing to perform the cointegration analysis of Johansen (1991) even in the presence of explosive behaviour. Particularly, we investigate the oil price run-up in the WTI crude oil market between July 2007 and July 2008 as this period is indicated as explosive as well as the oil price collapse between November 2015 and February 2016, to argue whether the 2007-2008 oil price run-up can be attributed to the existence of a speculative bubble and whether the oil price collapse exhibits any characteristics of bubble implosion. One of our main findings of this chapter is that contemporaneously, crude oil spot prices and all futures contracts contain both an explosive root and a unit root component from July 2007 to July 2008, whereas when we match the futures contract prices with the actual future spot prices then oil future prices of spot and the prices of the six month, twelve month and eighteen month futures contracts contain both an explosive root and a unit root component for this period.

Examining the 2014-2015 crude oil price collapse we argue that contemporaneously, crude oil spot prices and the one month futures contract and crude oil spot prices and the three month futures contract contain both an explosive root and a unit root component between November 2015 and February 2016, whereas matching the futures contract prices with the actual future spot prices results in a single explosive root between the future spot prices and the three month futures contract therefore the system contains both an explosive root and a unit root component during this period. Therefore, we argue that both oil prices of spot and futures contracts are $I(1,x)$ processes and the two variables cointegrate such that their linear combination is an $I(0)$ process for the periods July 2007 to July 2008 and November 2015 to February 2016. This is in support of the view commonly stated in the empirical literature that prices of spot and (short maturity) futures contracts should cointegrate even when there is a bubble episode in the sample (Engsted, 2006).[17]

As an extension to our study, we apply a date-stamping technique to the difference between the future spot prices and the futures contract prices as proposed by Pavlidis et al. (2017) that results in a delayed identification of the origination date of the bubble oil episode of 2007-2008 providing no statistical evidence of explosive behaviour between

[17] The notation $I(1,x)$ stands for variables with both explosive and random walk components and $I(x)$ for variables with just explosive common trends as in Nielsen (2010).

July 2007 and July 2008. Furthermore, applying the same date-stamping technique to the reverse series of the difference between the future spot prices and the futures contract prices results in a delayed identification of the origination date of the oil price collapse episode of 2014-2015 providing no statistical evidence of explosive behaviour (in the reverse series, therefore no market collapse in the original series as in Phillips and Shi 2018) between November 2015 and February 2016. These findings are consistent with our argument that during the peak of the oil price run-up of 2007-2008 and the oil price collapse of 2014-2015, crude oil future spot prices and futures contract prices are cointegrated, therefore their linear relationship is stationary and since the characteristic roots of their VAR model are, in some cases, explosive we conclude that oil prices of the spot and futures contracts coexplode during these two periods of interest.

The remainder of this chapter is organised as follows. Section 4.2 outlines the model and assumptions of testing for cointegration and co-explosiveness utilising the Johansen cointegration rank test and estimating a coexplosive vector autoregressive model. In Section 4.3 we present the Granger-Johansen representation theorem whereas in Section 4.4 we review how statistical analysis and hypothesis testing is conducted in a cointegration framework. Section 4.5 provides the limit theory around the Johansen cointegration test. Section 4.6 presents an empirical application on the WTI crude oil spot prices and futures contracts. Section 4.7 concludes. Tables and Figures are presented in sections 4.8 and 4.9 respectively.

In what follows, for a full column rank matrix α, we let $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$, while $\alpha_\perp$ denotes a basis to the orthogonal complement of span of α so $\alpha'_\perp \alpha = 0$ and $(\alpha, \alpha_\perp)$ is invertible. The notation a.s. P and D is used for properties holding almost surely, in probability and in distribution, respectively.

4.2 The Model and Assumptions

This section is structured as follows. We firstly present the simple vector autoregressive model as a starting point of our analysis. To continue with, we introduce the restriction of cointegration, followed by the restriction of co-explosiveness.

4.2.1 The VAR Model

Consider a p-dimensional time series vector X_t of order k. The vector autoregressive model is given by;

$$X_t = \mu + \theta t + \sum_{j=1}^{k} A_j X_{t-j} + \varepsilon_t \text{ for } t = 1, ..., T \tag{4.1}$$

where the innovation term $\varepsilon_t \sim^{iid} N_p(0, \Omega)$ is a martingale difference sequence, $A_j, \Omega \in \mathbb{R}^{p \times p}$, Ω is positive definite and $\mu, \theta \in \mathbb{R}^p$.

4.2.2 The Cointegrated VAR Model

Suppose a p-dimensional time series vector X_t of order k containing $I(1)$ variables as in Johansen (1995). Then, Equation (4.1) can be reparameterised in equilibrium correction form as;

$$\Delta_1 X_t = \mu + \Pi X_{t-1} + \Pi_l t + \sum_{j=1}^{k-1} \Gamma_j \Delta_1 X_{t-j} + \varepsilon_t \text{ for } t = 1, ..., T \tag{4.2}$$

where Δ_1 is a first difference operator defined as $\Delta_1 X_t = X_t - X_{t-1}$, the innovation term $\varepsilon_t \sim^{iid} N_p(0, \Omega)$, $\Gamma_j, \Omega \in \mathbb{R}^{p \times p}$, Ω is positive definite and $\mu \in \mathbb{R}^p$. The usual reduced-rank cointegration hypothesis applies;

$H_1(r): \; rank(\Pi, \Pi_l) \leq r$

and under the reduced rank restriction, Equation (4.2) can be written as;

$$\Delta_1 X_t = \mu + \alpha(\beta_1' X_{t-1} + \delta_1' t) + \sum_{j=1}^{k-1} \Gamma_j \Delta_1 X_{t-j} + \varepsilon_t \text{ for } t = 1, ..., T \tag{4.3}$$

where again Δ_1 is a first difference operator defined as $\Delta_1 X_t = X_t - X_{t-1}$, the innovation term $\varepsilon_t \sim^{iid} N_p(0, \Omega)$, $\Gamma_j, \Omega \in \mathbb{R}^{p \times p}$, Ω is positive definite, $\mu \in \mathbb{R}^p$, $\delta_1 \in \mathbb{R}^r$ and $\alpha, \beta_1 \in \mathbb{R}^{p \times r}$.

4.2.3 The Coexplosive VAR Model

Nielsen (2010) introduces the coexplosive VAR model to examine the presence of both a unit root and a single explosive root in the series. To avoid inconsistency problems that can arise with multiple explosive roots, Nielsen (2010) focuses on the case that there is just one positive explosive root in the system.

We now assume that the characteristic polynomial for Equation (4.3) has a single positive explosive root, $\rho > 1$. As a restriction to Equation (4.3), the following coexplosive model arises as suggested by Nielsen (2010);

$$\Delta_1 \Delta_\rho X_t = \mu + \Pi_1 \Delta_\rho X^*_{t-1} + \Pi_\rho \Delta_1 X_{t-1} + \sum_{j=1}^{k-2} \Phi_j \Delta_1 \Delta_\rho X_{t-j} + \varepsilon_t \tag{4.4}$$

where Δ_1 is a first difference operator defined as $\Delta_1 X_t = X_t - X_{t-1}$ and $\Delta_\rho X_t = X_t - \rho X_{t-1}$ is a ρ order difference operator and $\Delta_\rho X^*_{t-1} = \{\Delta_\rho X^{'}_{t-1}, (1-\rho)t\}^{'}$. The innovation term $\varepsilon_t \sim^{iid} N_p(0, \Omega)$, $\Pi_1, \Pi_\rho, \Phi_j \in \mathbb{R}^{p \times p}$, $\mu \in \mathbb{R}^p$ and $\rho \in \mathbb{R}$.

Under the hypothesis of reduced rank, Equation (4.4) can be rewritten as;

$$\Delta_1 \Delta_\rho X_t = \mu + \alpha_1 \beta_1^{*'} \Delta_\rho X^*_{t-1} + \alpha_\rho \beta^{'}_\rho \Delta_1 X_{t-1} + \sum_{j=1}^{k-2} \Phi_j \Delta_1 \Delta_\rho X_{t-j} + \varepsilon_t \tag{4.5}$$

where $\beta^*_1 = (\beta^{'}_1, \delta^{'}_1)^{'}$ Δ_1 and Δ_ρ are a first and ρ difference operator respectively. The new parameters α_1, α_ρ, β_ρ and Φ_j depend non-linearly on the original parameters in Equation (4.3);

$$\alpha_1 = \frac{\alpha}{1-\rho}, \qquad \alpha_\rho \beta^{'}_\rho = -\rho \left(I_p + \frac{a\beta^{/}_1}{1-\rho} - \sum_{j=1}^{k-1} \rho^{-j} \Gamma_j \right), \qquad \Phi_j = \sum_{l=j+1}^{k-1} \rho^{j-l} \Gamma_l \tag{4.6}$$

and the characteristic polynomial for Equation (4.3) or Equation (4.5) is given by the determinant of;

$$(1 - z^{-1}) I_p - z^{-1} \alpha \beta^{'}_1 - \sum_{j=1}^{k-1} z^{-j} (1 - z^{-1}) \Gamma_j = \frac{z-1}{z} \left(I_p + \frac{\alpha \beta^{'}_1}{1-z} - \sum_{j=1}^{k-1} z^{-j} \Gamma_j \right). \tag{4.7}$$

In case that $z = \rho$, Equation (4.7) reduces to $(1 - \rho^{-1}) \alpha_\rho \beta^{'}_\rho$, $rank(\alpha_\rho \beta^{'}_\rho) = p - 1$ and therefore there is one characteristic root at ρ. The new parameters vary freely, so $\alpha_1, \beta_1 \in \mathbb{R}^{p \times r}, \alpha_\rho, \beta_\rho \in \mathbb{R}^{p \times (p-1)}, \mu \in \mathbb{R}^p, \Phi_j, \Omega \in \mathbb{R}^{p \times p}$ so Ω is positive definite and $\rho > 1$.

4.3 The Granger-Johansen Representation

In order to interpret the parameters of Equation (4.5) we use *Assumption 4.3.*

Assumption 4.3. *The parameters need to satisfy the following conditions:*

1. *The matrices* $\alpha_1, \beta_1 \in \mathbb{R}^{p\times r}$ *and* $\alpha_\rho, \beta_\rho \in \mathbb{R}^{p\times(p-1)}$ *have full column rank.*
2. *If* $|A(z)| = 0$ *then* $|z| = 1$ *or* $|z| = \rho$ *where* $\rho > 1$ *meaning that the non-stationary characteristic roots of* X_t *are either at* 1 *or* ρ*, where* $\rho > 1$.
3. *The* $\det(\alpha_{1\perp}^{'}\Psi_1\beta_{1\perp}) \neq 0$ *and* $\det(\alpha_{\rho\perp}^{'}\Psi_\rho\beta_{\rho\perp}) \neq 0$ *where;*

$$\Psi_1 = I_p + \frac{\alpha_\rho\beta_\rho^{'}}{\rho-1} - \sum_{j=1}^{k-2}\Phi_j,\ \Psi_\rho = I_p + \frac{\alpha_1\beta_1^{'}}{1-\rho} - \sum_{j=1}^{k-2}\rho^{-j}\Phi_j.$$

where the parameters β_1 *and* β_ρ *reflect the cointegrating and coexplosive relationships.*

As noted by Nielsen (2010) and given *Assumption 4.3* the Granger's representation theorem (Engle and Granger 1987) can be formulated as follows.

Theorem 4.4. *Consider Equation* (4.5) *and suppose Assumption 4.3: condition 1 holds. Then;*

$$U_t = \{(\Delta_\rho X_t^*)^{'}\beta_1^*, (\Delta_1 X_t)^{'}\beta_\rho, (\Delta_1\Delta_\rho X_t)^{'}, ..., (\Delta_1\Delta_\rho X_{t-k+3})^{'}\}^{'}$$

can be given a stationary initial distribution ensuring the representation;

$$X_t \stackrel{\mathsf{D}}{=} \frac{1}{1-\rho}C_1\sum_{s=1}^{t}\varepsilon_s + \frac{1}{\rho-1}C_\rho\sum_{s=1}^{t}\rho^{t-s}\varepsilon_s + Y_t + \tau_c + \tau_\ell t + \tau_x\rho^t,$$

where $C_x = \beta_{x\perp}(\alpha_{x\perp}^{'}\Psi_x\beta_{x\perp})^{-1}\alpha_{x\perp}^{'}$ *and* Y_t *is a stationary process. In particular,* $\beta_c^{'}X_{t-1}$ *can be given a stationary initial distribution for any* $\beta_c \in span(\beta_1) \cap span(\beta_\rho)$.

The linear slope coefficient τ_l *can be defined as;*

$$\tau_l = \frac{C_{1\mu}}{1-\rho} + (C_1\Psi_1 - I(p))\bar{\beta}_1\delta_1^{'}$$

so $\beta_1^{'}\tau_l + \delta_1^{'} = 0$*. The coefficients for the exponential term* τ_x*, and the constant level* τ_c *depend on the initial values in such a way that* $\beta_\rho^{'}\tau_x = 0$ *and;*

$$\beta_1^{'}\tau_c = \bar{\alpha}_1^{'}\left(\frac{\Psi_1C_1 - I_p\mu}{1-\rho}\right) + \bar{\alpha}_1^{'}\left(\Psi_1C_1\Psi_1 - \Psi_1\right)\bar{\beta}_1\delta_1^{'} + \frac{\delta_\rho^{'}}{(1-\rho)}$$

Finally, $\tilde{X}_t = X_t - \tau_c - \tau_\ell t$ *satisfies the following equation;*

$$\Delta_1\Delta_\rho\tilde{X}_t = \alpha_1\beta_1^{'}\Delta_\rho\tilde{X}_{t-1} + \alpha_\rho\beta_\rho^{'}\Delta_1\tilde{X}_{t-1} + \sum_{j=1}^{k-2}\Phi_j\Delta_1\Delta_\rho\tilde{X}_{t-j} + \varepsilon_t. \tag{4.8}$$

As a consequence of the representation, under *Assumption 4.3, condition 1* the process that satisfies Equation (4.5) has $p-r$ random walk (unit root) components and one explosive root ρ.[18] The coefficients β_1 and β_ρ are vectors and can be interpreted as cointegrating and coexplosive relationships respectively, in that $\beta_1' X_{t-1}$ has no random walk component, while $\beta_\rho' X_{t-1}$ has no explosive trend.

4.4 Statistical Analysis and Hypothesis Testing

In order to determine the cointegrating rank of Equation (4.2), the reduced rank hypothesis $H_1(r) : rank(\Pi, \Pi_l) \leq r$ needs to be tested. The likelihood ratio test statistic is;

$$LR = -2\big((T-k)\ln L_T(\hat{\theta}_0) - (T-k)\ln L_T(\hat{\theta}_1)\big) \tag{4.9}$$

where $\hat{\theta}_0$ is the restricted parameter that corresponds to the reduced-rank model and $\hat{\theta}_1$ is the unrestricted parameter that corresponds to the full rank model, T is the sample size and k is the number of lags. Using the eigen decomposition form of the log-likelihood function;

$$\ln L_T(\hat{\theta}_0) = -\frac{T}{2}(1+\ln 2\pi) - \frac{1}{2}\ln|S_{00}| - \frac{1}{2}\sum_{i=1}^{r}\ln(1-\hat{\lambda}_i) \tag{4.10}$$

$$\ln L_T(\hat{\theta}_1) = -\frac{T}{2}(1+\ln 2\pi) - \frac{1}{2}\ln|S_{00}| - \frac{1}{2}\sum_{i=1}^{p}\ln(1-\hat{\lambda}_i). \tag{4.11}$$

The estimation of the parameters $\hat{\theta}_0$ and $\hat{\theta}_1$ happens as follows. We firstly estimate the residual vectors $R_{0,t}$ and $R_{1,t}$ from regressing $\Delta_1 X_t$ and $(X_{t-1}')'$ on $\Delta_1 X_{t-1}, ...,$ $\Delta_1 X_{t-k+1}$ and a constant and then we find the sum of square matrices

$$S_{ij} = T^{-1}\sum_{i=1}^{T}\hat{R}_{i,t}\hat{R}_{j,t}^{'} \;\; i,j = 0,1. \tag{4.12}$$

[18] The case of one explosive root is emphasised here.

To continue with, after computing the Choleski decomposition $S_{11} = LL'$ we perform an eigen decomposition on $L^{-1}S_{10}S_{00}^{-1}S_{01}L^{-1'}$ and we estimate the eigenvalues $\hat{\lambda}$ and matrix of eigenvectors E. Finally, we normalise the eigenvector matrix $L^{-1'}E$ to obtain the coefficients of Equation (4.3). The maximised log likelihood function of Equation (4.3) can now be written as;

$$\hat{\ell} = \max \ell(\theta) = -\frac{T}{2}\Big\{ \log\det(S_{00}) + \sum_{j=1}^{r} \log(1-\hat{\lambda}_j)\Big\} \tag{4.13}$$

where $S_{00} = T^{-1}\sum_{t=1}^{T} R_{0,t}R_{0,t}'$. We can now form an alternative likelihood ratio test statistic, known as the trace statistic given by;

$$LR\{H_1\} = -T\sum_{j=r+1}^{p} \log(1-\hat{\lambda}_j). \tag{4.14}$$

The maximum likelihood estimators can be easily found by taking advantage of the reparameterisation. Thus, the estimated explosive root of the characteristic polynomial for Equation (4.3) will be the estimated explosive root, ρ. The parameters of Equation (4.5) can be estimated using Equation (4.6).

By knowing the coexplosive vector ρ, we can maximise the log likelihood function. In particular, θ's can be estimated for a given value of ρ as outlined above; the residual vectors $R_{0,t}(\rho)$ and $R_{1,t}(\rho)$ are computed and then the eigenvalues and eigenvectors are estimated. The log likelihood function for ρ will be;

$$\hat{\ell}(\rho) = \max \ell(\rho,\theta) = -\frac{T}{2}\Big[\log\det\{S_{00}(\rho)\} + \sum_{j=1}^{r} \log\{1-\hat{\lambda}_j(\rho)\}\Big] \tag{4.15}$$

where $S_{00}(\rho) = T^{-1}\sum_{t=1}^{T}\{R_{0,t}(\rho)\}\{R_{0,t}(\rho)\}'$. The log likelihood function of Equation (4.15) can be maximised by running the test over all possible ρ's with $\rho > 1$. Finally, we estimate the likelihood ratio test statistic by testing against the hypothesis of full rank;

$$LR = -2(\hat{\ell}_0 - \hat{\ell}_1) \tag{4.16}$$

where $\hat{\ell}_0$ and $\hat{\ell}_1$ are the maximum likelihood estimators that correspond to the reduced-rank model and the full rank model respectively. According to Nielsen (2010) *Corollary 1* the likelihood ratio test statistic follows a $\chi^2_{r\times(p-r)}$ distribution with $r\times(p-r)$ degrees of freedom.

4.5 Limit Theory

As noted by Nielsen (2010), to derive the asymptotic distribution of the cointegration rank test, the following assumptions need to be made first given that $(\varepsilon_t, \mathscr{F})$ is a martingale difference sequence for some algebraic filtration;

Assumption 4.4. *For some* $\gamma > 0$ *exists, such that* $\sup_t E\{(\varepsilon_t'\varepsilon_t)^{(2+\gamma)/2}|\mathscr{F}_{t-1}\} < \infty$ *a.s.*

Assumption 4.5. *Suppose* $E\{(\varepsilon_t'\varepsilon_t)|\mathscr{F}_{t-1}\} = \Omega$ *a.s., where* Ω *is positive definite.*

Assumption 4.4 is utilised here to set an upper bound to the fluctuations of the error term whereas Assumption 4.5 makes the conditional heteroskedasticity time-invariant. Johansen (1995) derived the limit theory of the cointegration rank test under the assumption that the number of the unit roots is $p - r$ and the remaining characteristic roots are stationary.[19]

Theorem 4.5. *As in Nielsen (2010), assume Assumptions 4.1-4.3 hold and suppose model* (4.5)*. The asymptotic distribution of* (4.14) *will be given as in Johansen (1995) by;*

$$LR\{H_1\} \xrightarrow{D} tr\left\{ \int_0^1 dB_u F_u' \left(\int_0^1 F_u F_u' du \right)^{-1} \int_0^1 F_u dB_u' \right\}, \tag{4.17}$$

where $F_u = (B_u' - \int_0^1 B_s' ds, u - 1/2)'$ *with* B_u *being a* $p - r$ *standard Brownian motion.*[20]

As mentioned earlier, the likelihood ratio test statistic is asymptotically distributed as χ^2 with $r \times (p - r)$ the degrees of freedom (Nielsen 2010).

Corollary 4.1. *As noted by Nielsen (2010), under the assumption that the coexplosive vectors in* (4.5) *are known (*$H_x : \beta_\rho = \beta_\rho^o$*), suppose Equation* (4.5) *with* $\rho \geq \varrho$ *for* $\varrho > 1$*. Assume Assumptions 3.1-3.3 hold and that* $\tau'\varepsilon_t \sim^{iid} N(0, \tau'\Omega\tau)$*. Then;*

$$LR(H_x) \xrightarrow{D} \chi^2_{r\times(p-r)}$$

[19] According to Nielsen (2010) the last assumption is not necessary.
[20] For the case of no deterministic constant or trend see Johansen (1995) Chapter 6, Theorem 6.1.

where $\tau = (I_2 - \tau_\perp \bar{\tau}_\perp^{'})\alpha_\rho$, $\tau_\perp = \Psi_\rho \beta_{\rho\perp}$ *a non-zero vector due to Assumption 4.1 that formulates the asymptotic result and* $r \times (p - r)$ *the degrees of freedom.*

The asymptotic theory presented above is consistent with rational expectations since the distribution of the likelihood ratio test has been derived under the assumption that the innovation term, ε_t is a $\mathscr{F}-$ martingale difference sequence.

4.6 Empirical Application

This section discusses an empirical application of the cointegration rank test of Johansen (1988) as discussed by Nielsen (2010).

4.6.1 Data

We download WTI crude oil prices from Eikon for the period September 1995 to July 2019 and construct monthly and weekly series for each spot and associated futures contracts. In particular, our dataset contains crude oil prices for spot and nine futures contracts on NYMEX from 1995 to 2019 at weekly frequency. The futures contract maturities are one month, three, six, nine, twelve, fifteen, eighteen, twenty one, and twenty four months. The NYMEX crude oil future contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month.

WTI crude oil spot prices together with the percentage change of WTI crude oil price on a year earlier can be seen in Figure 22. The oil prices move upward from the beginning of 2007 until mid-2008. In particular, crude oil prices rise steadily until the end of 2007, followed by explosive growth during the first six months of 2008. In half a year, prices increased more than 50% of their nominal value to collapse during the second half of 2008 reverting back to 2004 prices (Saporta, Tudela and Trott 2009).

4.6.2 Cointegration Tests

Testing for cointegration across the whole sample period, that is September 1995 to July 2019, cannot give us information about any potential breaks to the cointegrating relationship that can be explained by price collapses as the unit root component might dominate the explosive component, leaving no evidence of cointegration breaks. For this reason, we perform the Johansen cointegration analysis recursively across different subsamples for the period September 1995 to July 2019 at 5% level of significance. We fix the end of the sample at the date that the WTI crude oil spot price reached its highest value for the 2007-2008 oil price bubble episode whereas focusing on the 2014-2015 oil price collapse, we fix the end of the subsample at the date that the WTI crude oil spot price dropped to its lowest level whereas we let the beginning of the subsample to change recursively, subject to a minimum window size that allows us to perform the Johansen cointegration test (ten observations). A constant is included in the cointegrating VAR model but no trend and the lag length is set to three, following Nielsen (2010).

We see that from June 2004, when the oil price run-up started, to July 2008 when the oil price collapsed, crude oil spot prices are cointegrated (recursively) with all future contracts prices except the three month one (see Figure 23). Additionally, crude oil spot prices are cointegrated between November 2015 and February 2016 marking the period of the 2014-2015 crude oil price collapse.

We consider the periods from July 2007 to July 2008 and November 2015 to February 2016 as our subsamples since the former contains the peak of the 2008 bubble episode, whereas the latter contains the 2014-2015 oil price collapse, ignoring any short-term blips. During the second subsample, from November 2015 to February 2016, both WTI crude oil spot and futures contract prices exhibit a severe price decline. Bearing in mind that expansion and collapse are the two main aspects of an asset bubble episode whereas market collapse can be as much of importance and impact as market expansion, we want to examine whether a coexplosive relationship still holds when an asset price series collapses. We choose to arrange the series of interest $\{y_t\}$ in reverse order such that $y_t^* = y_{T+1-t}$ for $t = 1, 2, ..., T$, from the period November 2015 to February 2016 in line with Philips and Shi (2018). Expanding the subsample to capture longer periods of explosiveness would result in rather inconclusive statistical inference as unit root and cointegration tests might not be able to identify bubble episodes that continuously grow and burst over time in the case the unit root component dominates the explosive component (Evans 1991).

4.6.3 Coexplosiveness: the contemporaneous case

Conducting a multivariate analysis on the logarithm of prices for the spot and futures contracts we estimate a bivariate VAR model for the sample periods July 2007 to July 2008 and November 2015 to February 2016 at weekly frequency, fitting three lags and including an intercept in the cointegrating regression. According to *Theorem 3.2*, the likelihood ratio test has the same asymptotic distribution as in Johansen (1995, Chap. 6) even in the presence of an explosive root.

Table 7 reports the cointegration rank tests for the period July 2007 to July 2008. The likelihood ratio test suggests that the cointegrated VAR model has a rank of one for the crude oil spot prices and all futures contracts. The trace test-statistic ranges from 21.66 for the fifteen month futures contract to 24.08 for the three month futures contract. All trace test statistics are larger than the critical value of 20.26 and therefore we reject the null hypothesis of zero rank at 5% level of significance. Additionally, under the alternative hypothesis of full rank, the trace test-statistic is smaller than the 9.16 critical value at 5% significance level for all futures contracts and therefore we do not reject the null hypothesis of reduced rank of one. Similarly, in Table 9 we perform the likelihood ratio test over the period November 2015 to February 2016. The trace test-statistic is greater than the critical value at 5% across all maturities of the futures contracts.

For all maturities of the futures contracts, the trace test-statistics exceed the 5% critical values and as the p-values are less than 5%, we can reject the null hypothesis of zero rank of the cointegrated VAR model, therefore there is statistical evidence of cointegration in the explosive subsample of our interest. The crude oil spot prices are cointegrated with the prices of the one, three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four futures contract during the periods July 2007 - July 2008 and November 2015 - February 2016.

4.6.4 Coexplosiveness: the non-contemporaneous case

So far we have made use of contemporaneous crude oil prices for spot and futures contracts. The crude oil spot price on a particular date is combined with the price of the futures contract (of a particular maturity) on the same date. Additionally, instead of using contemporaneous crude oil spot prices and futures contracts we perform the cointegration analysis on the actual future spot prices (s_{t+n}) and futures contract prices ($f_{t,n}$). In other words, we match the spot price with the futures contract price on the expiration date. For instance, we combine the crude oil spot price today with the price of the one month futures contract that expires today.

Following the same methodology, in Tables 8 and 10 we report the log-likelihood and trace test-statistics of the cointegration rank test by estimating a bivariate VAR model between crude oil future spot prices and the futures contract prices of the nine contracts across two subsample periods, July 2007 to July 2008 and November 2015 to February 2016 at a weekly frequency, fitting three lags and an intercept in the cointegrating regression.

In Table 8, the trace test-statistics suggest that the cointegrated VAR model has a reduced rank of one for the crude oil prices of the spot and one, six, twelve and eighteen month futures contract with a p-value less than 5% for the one, six and twelve month futures contract and less than 10% for the eighteen months futures contract rejecting the null of zero rank in favour of the alternative hypothesis of at least full rank for all four contracts. To continue, under the null of reduced rank of one the trace test-statistic does not exceed the 5% critical value of 9.16 and therefore we cannot reject the null hypothesis of rank one for the first three contracts (i.e. one, six and twelve month futures contract) and the 10% critical value of 7.56 for the eighteen month futures contract to conclude that the crude oil prices of the spot and the one, six, twelve, eighteen months futures contract are cointegrated from July 2007 to July 2008.

In Table 10, the trace test-statistics indicate that the bivariate cointegrated VAR model has a reduced rank of one when the crude oil spot price is considered alongside the one, three, six, twelve and eighteen month futures contract with a p-value less than 5% for all five futures contracts rejecting the null of zero rank in favour of the alternative hypothesis of at least rank one. In particular, under the null of reduced rank of one the trace test-statistic does not exceed the 5% critical value of 9.16 and therefore we cannot reject the null hypothesis of rank one for the one, three, six, twelve and eighteen month futures contract, concluding that the crude oil prices of the spot and the one, three, six, twelve, eighteen months futures contract are cointegrated for the period November 2015 to February 2016.

To continue, a VAR model as Equation (4.1) with an intercept and three lags is fitted. Table 11 presents the estimated roots of the characteristic polynomial for the periods July 2007 to July 2008 and November 2015 to February 2016 both for the contemporaneous and non-contemporaneous series. In the contemporaneous series, the WTI crude oil spot prices match with the futures contract prices on the same date whereas in the non-contemporaneous case, the WTI crude oil future contract prices match with the futures contract prices on the expiration date of the futures contract. As it can been seen in Table 11 for the period July 2007 to July 2008, the vector autoregression for the crude oil spot prices and the one month futures contract has a characteristic root of 1.0107 that is larger than one, indicating an explosive root in the system. That applies to the longer maturity contracts as well, since their vector autoregressive systems appear to have explosive characteristic roots that range between 1.016 for the three month futures contract to 1.029 for the twenty one month futures contract. In Table 11 we see that for the period November 2015 to February 2016 and when the series are contemporaneous, the vector autoregression for the crude oil spot prices and the one and three month futures contract have a characteristic root of 1.035 and 1.057 respectively that are larger than one, indicating an explosive root in the two systems.

From Table 7, we have inferred that there is evidence of cointegration between the crude oil spot and the one, three, six, eight, twelve, fifteen, eighteen, twenty one and twenty four months futures contract prices and since the estimated roots of the characteristic polynomials for all futures contracts are explosive, we conclude that there is statistical evidence of co-explosiveness between the crude oil prices of the spot and futures contracts between July 2007 and July 2008.

Furthermore, from Table 9 we have concluded that there is evidence of cointegration between the crude oil spot and the one, three, six, eight, twelve, fifteen, eighteen, twenty one and twenty four months futures contract prices and as the characteristic roots of the VAR model described above for the one and three month futures contract are explosive, we infer that there is statistical evidence of co-implosiveness between the crude oil prices of the spot and the one month and three month futures contract between November 2015 and February 2016.

In Table 11 we present the estimated roots of the characteristic polynomial across all futures contracts for the period July 2007 to July 2008 and November 2015 to February 2016 both at a weekly frequency when the series are non-contemporaneous. In Table 11 we see that all characteristic roots for the crude oil future prices of spot and all futures contracts, with the exception of the one month and three month futures contacts, are larger than one and therefore explosive. As in the contemporaneous case, the oil prices of the spot and futures contracts coexplode during the subsample July 2007 to July 2008, if they are cointegrated and the VAR model contains an explosive root. Combining these two, we can infer that there is statistical evidence of co-explosiveness between the crude oil spot prices and the six month, twelve month (both at 5%) and eighteen month (at 10%) futures contract.

As we can see in Table 11, the characteristic roots for the crude oil future prices of spot and the three, nine and twenty one month futures contract are larger than one and therefore explosive for the period November 2015 to 2016. Again, as in the contemporaneous case, the oil prices of the spot and futures contracts co-implode during the subsample period November 2015 to February 2016, if they are cointegrated and the VAR model contains an explosive root. Bearing this in mind, we can conclude that there is statistical evidence of co-implosiveness between the crude oil spot prices and the three month futures contract between November 2016 and February 2016 when we match the future spot price with the futures contract price on the expiration date.

4.6.5 Causality Testing

A question arising after finding evidence of cointegration between crude oil spot and futures contract prices, is what is the direction of the casual impact of the one series on the other and thus we test for Granger causality. Consider a bivariate VAR(2) model with a constant:

$$\begin{bmatrix} r_{s_t} \\ r_{f_{t,n}} \end{bmatrix} = \begin{bmatrix} \Phi_{s,0} \\ \Phi_{f,0} \end{bmatrix} + \begin{bmatrix} \Phi_{ss,1} & \Phi_{sf,1} \\ \Phi_{fs,1} & \Phi_{ff,1} \end{bmatrix} \begin{bmatrix} r_{s_{t-1}} \\ r_{f_{t-1,n}} \end{bmatrix} + \begin{bmatrix} \Phi_{ss,2} & \Phi_{sf,2} \\ \Phi_{fs,2} & \Phi_{ff,2} \end{bmatrix} \begin{bmatrix} r_{s_{t-2}} \\ r_{f_{t-2,n}} \end{bmatrix} + \begin{bmatrix} \Phi_{ss,3} & \Phi_{sf,3} \\ \Phi_{fs,3} & \Phi_{ff,3} \end{bmatrix} \begin{bmatrix} r_{s_{t-3}} \\ r_{f_{t-3,n}} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s,t} \\ \varepsilon_{f,n,t} \end{bmatrix} \tag{4.18}$$

where r_{s_t} are the returns on WTI crude oil spot prices, $r_{f_{t,n}}$ the returns on WTI crude oil futures contracts, n is the futures contract length and $[\varepsilon_{s,t} \quad \varepsilon_{f,n,t}]'$ the heteroskedastic error term vector. We make use of the likelihood ratio test to investigate whether it is returns on spot prices (r_{s_t}) causing returns on futures contract prices ($r_{f_{t,n}}$ where n is the futures contract length) or vice versa. We fit three lags of spot and futures returns and we account for the case that the errors are heteroskedastic but uncorrelated by performing a heteroskedasticity robust likelihood ratio test. We test the hull hypothesis of no Granger causality across all possible restrictions, therefore:

$$H_{0,ss}: \quad \Phi_{ss,1} = \Phi_{ss,2} = \Phi_{ss,3} = 0$$

$$H_{0,sf}: \quad \Phi_{sf,1} = \Phi_{sf,2} = \Phi_{sf,3} = 0$$

$$H_{0,fs}: \quad \Phi_{fs,1} = \Phi_{fs,2} = \Phi_{fs,3} = 0$$

$$H_{0,ff}: \quad \Phi_{ff,1} = \Phi_{ff,2} = \Phi_{ff,3} = 0.$$

According to the null hypothesis $H_{0,ss}$, lag of returns on WTI crude oil spot prices do not Granger cause returns on WTI crude oil spot prices at time t. Respectively, according to the null hypothesis $H_{0,sf}$, lag of returns on WTI crude oil futures contract prices (of length n) do not Granger cause returns on WTI crude oil spot prices at time t whereas not rejecting the null hypothesis $H_{0,fs}$, would mean that lag of returns on WTI crude oil spot prices do not Granger cause returns on WTI crude oil futures contract prices (of length n) at time t. Finally, under to the null hypothesis $H_{0,ff}$, lag of returns on WTI crude oil futures contract prices (of length n) do not Granger cause returns on WTI crude oil futures contract prices (of length n) at time t.

In Table 12, we report the probabilities of rejection of the null hypotheses of no Granger causality between WTI crude oil returns on spot prices r_{s_t} and WTI crude oil returns on futures contract prices $r_{f_{t,n}}$. We find that lags of returns on short (one to three months) and mid maturity WTI crude oil futures contracts (six to nine months at 5% level of significance) Granger-cause returns on spot prices for the entire sample period September 1995 to July 2019. Furthermore, we see that lags of returns on short (one to three months) and mid maturity WTI crude oil futures contract prices (six months at 5% level of significance) Granger-cause returns on WTI crude oil futures contract prices for the same period. Finally, there is some Granger causality between lags of returns on long maturity (eighteen to twenty four months) crude oil futures contracts and returns on spot prices for the entire sample period September 1995 to July 2019.

To summarise, in the contemporaneous case crude oil spot prices and all futures contracts contain both an explosive root and a unit root component whereas when we match the futures contracts prices ($f_{t,n}$) with the actual future spot prices (s_{t+n}) then crude oil future spot prices and the six month, twelve month and eighteen month futures contract prices contain both an explosive root and a unit root component for the period

July 2007 to July 2008.

4.6.6 Unit Root Testing for Rational Bubbles

Interested in investigating whether a coexplosive relationship still holds when a price series collapses as the impact of a market collapse can be severe and therefore it is important to be able to identify crisis episodes, we run the BSADF test on the reverse series of the difference between future spot and futures contract prices.

In the rational bubbles literature, the difference between spot and futures contract prices is commonly perceived as a relationship between prices and fundamentals. Therefore, any deviations between future spot prices and futures contract prices for a prolonged period of time might be considered as deviations from fundamentals supporting the argument in favour of a bubble episode. However, in commodity markets it is not always obvious what is the fundamental of a particular commodity. In particular, in the oil market the convenience yield (i.e. costs of storage) is used as a proxy of market fundamentals (see for instance Tsvetanov et al. 2016). In the framework of this chapter, we consider the fluctuations between future spot and futures contract prices as evidence of market inefficiency (rather than deviations from fundamentals) that could, potentially, lead to a bubble episode.

As a robustness check, we investigate the oil price run-up of 2007-2008 as well as the oil price collapse of 2014-2015 further. We compare our results to Pavlidis et al. (2017), according to which it is the fundamentals that are responsible for the fluctuations in oil prices in the early 2000s and not a speculative bubble as the fundamental price of oil is practically unobservable and therefore any statistical inference in favour of a speculative bubble might be due to misspecification error. Finding no evidence of a bubble episode in the difference between future spot and futures contract prices using Pavlidis et al. (2017) approach during the periods we identified earlier as coexplosive supports our evidence that the series coexplode (or co-implode for the reverse series) during the oil price run-up of 2007-2008 and the oil price collapse of 2014-2015.

In particular, we apply the ADF, SADF and GSADF tests to the difference between the future spot prices (s_{t+n}) and futures contract prices ($f_{t,n}$) where n is the contract length, across all different maturity contracts for the sample period September 1995 to July 2019.

Table 13 summarises the unit root test results in differences together with the finite sample critical values. Both the ADF and SADF test-statistics are below the 95% critical values across all differences between future spot prices and futures contract prices failing to reject the null hypothesis of a unit root. The poor performance of the ADF and SADF tests is expected as the former one suffers from low power whereas the latter one under performs in a multiple bubble environment (Phillips et al. 2015).

In contrast, the GSADF test-statistics reject the null hypothesis of a unit root in favour of the alternative hypothesis of explosiveness, for all the differences between the spot prices and the mid and longer maturity futures contract prices, namely the differences between the six month future spot price and the six month futures contract price ($s_{t+6} - f_{t,6}$), the nine month future spot price and the nine month futures contract price ($s_{t+9} - f_{t,9}$), the twelve month future spot price and the twelve month futures contract price ($s_{t+12} - f_{t,12}$), the fifteen month future spot price and the fifteen month futures contract price ($s_{t+15} - f_{t,15}$), the eighteen month future spot price and the eighteen month futures contract price ($s_{t+18} - f_{t,18}$), the twenty one month future spot price and the twenty one month futures contract price ($s_{t+21} - f_{t,21}$) and the twenty four month future spot price and the twenty four month futures contract price ($s_{t+24} - f_{t,24}$).

To continue, in order to identify the exact periods of explosive behaviour in the oil market we apply the BSADF test of Phillips et al. (2015). Figures 24 to 26 plot the BSADF test-statistics sequence of the differences between the future spot price and the futures contract price for all nine futures contracts (i.e. $s_{t+n} - f_{t,n}$) together with their simulated finite sample critical value sequence at 95% level of significance. In addition, we apply the BSADF test in the reverse order of the series that represents the difference between the future spot price and the futures contract price across futures contracts with different maturities so that if $y_t = s_{t+n} - f_{t,n}$ where n is the contract length then $y_t^* = y_{T+1-t}$ for $t = 1, 2, ..., T$. Figures 27-29 illustrate the BSADF test-statistics sequence of the reversed series that represents the difference between the future spot price and the futures contract price across different contract lengths with the corresponding simulated finite sample critical value sequence at 95% level of significance.

The reverse regression approach is capable of identifying the origination date of the collapse as well as the termination date that corresponds to the market recovery date when either single or multiple market crashes take place since by reverse transformation, the original (mildly integrated) collapse process transforms to a (mildly) explosive process and vice versa (Philips and Shi 2018). Thus, the reverse regression approach is able to identify the market collapse, however it does so by utilising ex post data and therefore can be used for identifying historical episodes of crashes rather than real time bubble episodes.

Ignoring any short term blips and focusing on the oil crisis of 2007-2008 and the oil price collapse of 2014-2015, the BSADF test-statistics increase above the 95% critical value sequence for all the differences between the future spot prices (s_{t+n}) and futures contract prices ($f_{t,n}$) except for the one month futures contract and then drops again below the 95% critical value sequence.

To be more precise on the exact time period of the oil price run-up and collapse episode, we perform a date-stamping technique on the BSADF test-statistics calculated on the difference between the future spot price and the futures contract price for all contracts (i.e. $s_{t+n} - f_{t,n}$). Particularly, we define the origination date of the bubble episode as the first chronological observation of which the test statistic is greater than the simulated finite sample critical value therefore rejecting the null hypothesis of stationarity in favour of the alternative hypothesis of an end-of-sample bubble. On the same framework, the termination date of the bubble episode is defined as the first chronological observation of which the test-statistic becomes smaller than the simulated finite sample critical value. Respectively, the origination date of the collapse episode is given by the first chronological observation of which the test statistic of the reverse regression is greater than the simulated finite sample critical value whereas the termination date of the collapse episode and, therefore the market recovery date, is given by the first chronological observation of which the test-statistic becomes smaller than the simulated finite sample critical value.

Table 14 presents the periods of explosive episodes (origination and termination dates) for the period September 2015 to July 2019 as identified by the Phillips et al. (2015) methodology. Particularly, we apply BSADF test on the difference between the future spot price and the futures contract price for all nine contracts ($s_{t+n} - f_{t,n}$, where n is the length of the contract), and we compare these origination and termination dates against the respective dates of the bubble episodes that are suggested by applying the BSADF test on the crude oil future spot and futures contract prices separately. In Table

15, we report the market crash and recovery dates for the 2014-2015 oil price collapse by applying the BSADF test on the reverse series that represents the difference between crude oil future spot and futures contract prices as well as the the crude oil spot and futures contract prices across different contract lengths.

As we see in Table 14, when applied to the series separately, the BSADF test succeeds to detect the explosive episode of 2007-2008 on time, however when applied on the difference between crude oil future spot prices and futures contract prices the date-stamping results in delayed identification of the 2007-2008 bubble episode as the WTI crude oil spot price collapsed at the end of July 2008. Concerning the 2014-2015 oil price collapse, it can be seen that in Table 15 the BSADF test identifies the collapse date between January 2016 and March 2016 which is in line with the oil price collapse that took place in February 2016 whereas the market recovery date is set between March 2016 and May 2016.

Once again, applying the reverse regression BSADF test on the difference between the future spot price and the futures contract price across all contracts with different maturity results in either no identification of the crash episode at all or in a delayed identification as indicated by the differences between the nine month future spot price and the nine month futures contract price ($s_{t+9}-f_{t,9}$), the fifteen month future spot price and the fifteen month futures contract price ($s_{t+15}-f_{t,15}$), the eighteen month future spot price and the eighteen month futures contract price ($s_{t+18}-f_{t,18}$), the twenty one month future spot price and the twenty one month futures contract price ($s_{t+21}-f_{t,21}$) and the twenty four month future spot price and the twenty four month futures contract price ($s_{t+24}-f_{t,24}$).

Overall, in Tables 14 and 15 we can see that date-stamping the test-statistic of the BSADF test applied on the difference between the future spot price and the futures contract price results in a delayed identification of the origination date of the oil bubble episode of 2007-2008 and a delayed identification of the 2014-2015 crude oil price collapse. Pavlidis et al. (2017) methodology seems to identify the origination date of the 2007-2008 explosive episode as well as the origination date of the 2014-2015 oil price collapse episode with delay (or even not at all for some futures contracts) comparing to what the application of the BSADF test on the individual series suggests.

Particularly, the BSADF test on the three month future contract suggests the origination date of the oil bubble episode of 2007-2008 eight months earlier than the BSADF test on the difference between the future spot price in three months and the three month futures contract price ($s_{t+3} - f_{t,3}$). To continue, applying the BSADF test to mid maturity contracts (six to eighteen months) we identify the origination date of the bubble episode one year earlier compared to Pavlidis et al. (2017) (i.e. the difference between the future spot prices and the futures contract prices) whereas applying the BSADF test to longer-maturity contracts (twenty one to twenty four months) suggests that the origination date of the oil bubble episode more than four years earlier compared to Pavlidis et al. (2017) (i.e. the BSADF test on the difference between the future spot price and the futures contract price, $s_{t+n} - f_{t,n}$, where n is the length of the contract). Furthermore, applying the BSADF test on the differences between future spot and short-term length futures contract results in no identification of the 2014-2015 oil price collapse episode at all, whereas for mid and longer-maturity contracts the reverse regression BSADF test identifies the origination date of the collapse episode between between ten to thirteen months later compared to the application of the BSADF test on the spot and each futures contract prices separately. To conclude, there seems to be a significant delay in identifying the beginning of the oil bubble episode of 2007-2008 and the oil price collapse of 2014-2015 compared to the date-stamping results of the BSADF on the actual spot and future contract prices series.

Our cointegration analysis seems to be in accordance to Pavlidis et al. (2017) since applying the BSADF test on the future spot prices and the futures contract prices provide no statistical evidence of explosive behaviour on the differences between the future spot price and the futures contract price between July 2007 and July 2008 that we identify co-explosiveness and between November 2015 and February 2016 that we identify co-explosiveness at the reverse series and therefore co-implosiveness in the series itself, implying that the two linear relationships are a stationary process, as date-stamping identifies the origination date of the bubble and collapse episode on $s_{t+n} - f_{t,n}$, where n is the length of the contract, with delay across futures contracts with different maturities.

4.7 Conclusion

This chapter applies an extension of Johansen's cointegration rank test (Johansen 1988) that allows for explosive roots as suggested by Nielsen (2010). This approach can offer valuable information since it allows to test for cointegrating relationships across different series while simultaneously testing whether the series contain any explosive components, allowing to perform the cointegration analysis of Johansen (1991) even in the presence of explosive behaviour in the related series.

We have utilised Johansen's cointegration rank test to analyse an explosive episode in the WTI crude oil market between July 2007 and July 2008 and the oil price collapse between November 2015 and February 2016 at a weekly frequency as the former period is indicated as explosive and the latter one has characteristics of a market crash.

We provide evidence that during the period July 2007 to July 2008 when we choose to use contemporaneous crude oil prices for spot and futures contracts there is a single explosive root in the cointegrated VAR model between the crude oil spot prices and all futures contracts while at the same time the series are cointegrated, whereas when we match the actual future spot price with the futures contract price, there is a single explosive root in the cointegrated VAR model between the crude oil spot prices and the six month, twelve month and eighteen month futures contract while at the same time the series are cointegrated for that time period. Therefore, the series of the crude oil spot prices and futures contracts coexplode and their linear relationship is stationary between July 2007 and July 2008.

When we test for the oil price collapse between November 2015 and February 2016 contemporaneously, we find a single explosive root in the cointegrated VAR model between the crude oil spot prices and the one month futures contract and between the reverse series of the crude oil spot prices and the three month futures contract while at the same time the series are cointegrated. Matching the actual future spot price with the futures contract price, we conclude that there is a single explosive root in the cointegrated VAR model between the crude oil spot prices and the three month futures contract while at the same time the two series are cointegrated for that time period. Therefore, the series of the crude oil spot prices and futures contract co-implode and their linear relationship is stationary from November 2015 to December 2016.

Our findings suggest that both oil prices of spot and futures contracts are $I(1, x)$ processes and the two variables cointegrate such that their linear combination is an $I(0)$ process for the periods July 2007 to July 2008 and November 2016 to February 2017. This is in support of the view stated in the empirical literature that prices of spot and (short maturity) futures contracts should cointegrate even when there is a bubble episode in the sample (Engsted 2006).

Applying a date-stamping technique to the difference between the future spot prices and the futures contract prices results in a delayed identification of the origination date of the bubble oil episode of 2007-2008 and the oil price collapse of 2014-2015. This provides no statistical evidence of explosive behaviour between July 2007 and July 2008 as well as no statistical evidence of explosiveness of the reversed series between November 2015 and February 2016 in support of our evidence that during the oil run-up of 2007-2008 and the oil price collapse of 2014-2015, crude oil future spot prices and futures contract prices are cointegrated. Therefore their linear relationship is stationary and since the characteristic roots of their VAR model are, in most of the cases, explosive we conclude that oil prices of the spot and futures contracts coexplode/co-implode during these periods.

4.8 Tables

Table 7: Cointegration rank determination for the period July 2007 to July 2008, contemporaneous series

Hypothesis	**1 month**			**3 month**			**6 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	285.89	22.27	0.026	264.77	24.08	0.014	247.80	23.75	0.016
rank$\leq$ 1	294.34	5.38	0.276	273.88	5.85	0.205	256.32	6.71	0.143
rank$\leq$ 2	297.27			276.81			259.68		

Hypothesis	**9 month**			**12 month**			**15 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	237.56	22.66	0.0231	229.91	21.91	0.030	223.62	21.66	0.032
rank$\leq$1	245.52	6.74	0.141	237.48	6.76	0.140	231.03	6.84	0.136
rank$\leq$2	248.89			240.86			234.45		

Hypothesis	**18 month**			**21 month**			**24 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	218.41	21.78	0.031	214.72	22.02	0.029	211.59	22.13	0.028
rank$\leq$1	225.81	6.98	0.128	222.20	7.07	0.123	219.06	7.18	0.118
rank$\leq$2	229.30			225.73			222.65		

Note: Cointegration rank determination with an intercept restricted in the cointegrating regression. Critical values are based on Johansen (1995, Table 2) and Doornik (1998) and the tests are performed at 5% level of significance. The Table summarises results from estimating a cointegrated VAR model for the logarithmic prices for the spot and futures contracts across different maturities we estimate a bivariate VAR model for the sample period July 2007 to July 2008 at weekly frequency, fitting three lags and including an intercept in the cointegrating regression. The likelihood ratio test has the same asymptotic distribution as in Johansen (1995, Chap.6) even in the presence of an explosive root.

Table 8: Cointegration rank determination for the period July 2007 to July 2008, non-contemporaneous series (lags)

Hypothesis	**1 month**			**3 month**			**6 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	172.22	40.01	0.001	169.97	20.26	0.300	148.64	20.49	0.046
rank≤1	189.08	6.28	0.170	178.89	9.16	0.053	156.36	5.05	0.325
rank≤2	192.22			180.24			158.89		
Hypothesis	**9 month**			**12 month**			**15 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	139.76	14.22	0.305	139.46	20.90	0.041	144.04	22.98	0.021
rank≤1	146.30	1.14	0.932	145.43	8.96	0.055	150.11	10.84	0.024
rank≤2	146.87			149.91			155.53		
Hypothesis	**18 month**			**21 month**			**24 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	176.87	19.58	0.062	187.35	25.13	0.010	186.97	25.84	0.008
rank=≤1	183.05	7.22	0.116	194.63	10.57	0.027	193.61	12.57	0.011
rank≤2	186.66			199.91			199.89		

Note: Cointegration rank determination with an intercept restricted in the cointegrating regression. Critical values are based on Johansen (1995, Table 2) and Doornik (1998) and the tests are performed at 5% level of significance. The Table summarises results from estimating a cointegrated VAR model for the logarithmic prices for the future spot and futures contracts across different maturities we estimate a bivariate VAR model for the sample period July 2007 to July 2008 at weekly frequency, fitting three lags and including an intercept in the cointegrating regression. The likelihood ratio test has the same asymptotic distribution as in Johansen (1995, Chap.6) even in the presence of an explosive root.

Table 9: Cointegration rank determination for the period November 2015 to February 2016, contemporaneous series

Hypothesis	**1 month**			**3 month**			**6 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	40.32	22.80	0.022	38.10	21.57	0.033	37.45	20.97	0.040
rank≤1	49.15	5.14	0.311	45.45	6.86	0.134	45.15	5.57	0.246
rank≤2	51.72			48.88			47.93		

Hypothesis	**9 month**			**12 month**			**15 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	38.01	22.01	0.029	38.25	24.44	0.013	39.03	29.45	0.003
rank≤1	46.50	5.02	0.329	47.94	5.05	0.325	51.12	5.27	0.292
rank≤2	49.01			50.47			53.76		

Hypothesis	**18 month**			**21 month**			**24 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	33.37	33.49	0.001	39.68	37.30	0.001	39.59	40.36	0.001
rank≤1	53.33	5.56	0.249	55.41	5.84	0.207	56.72	6.10	0.184
rank≤2	56.11			58.33			59.77		

Note: Cointegration rank determination with an intercept restricted in the cointegrating regression. Critical values are based on Johansen (1995, Table 2) and Doornik (1998) and the tests are performed at 5% level of significance. The Table summarises results from estimating a cointegrated VAR model for the logarithmic prices for the spot and futures contracts across different maturities we estimate a bivariate VAR model for the sample period November 2015 to February 2016 at weekly frequency, fitting three lags and including an intercept in the cointegrating regression. The likelihood ratio test has the same asymptotic distribution as in Johansen (1995, Chap.6) even in the presence of an explosive root.

Table 10: Cointegration rank determination for the period November 2015 to February 2016, non-contemporaneous series (lags)

Hypothesis	**1 month**			**3 month**			**6 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	31.49	23.48	0.018	32.79	21.00	0.040	35.07	20.49	0.001
rank≤1	40.68	5.10	0.318	41.36	3.86	0.504	48.14	5.05	0.316
rank≤2	43.23			43.29			50.70		

Hypothesis	**9 month**			**12 month**			**15 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	33.78	14.24	0.304	35.52	64.18	0.041	43.24	15.44	0.203
rank≤1	38.51	4.79	0.364	58.95	17.32	0.055	48.26	5.41	0.271
rank≤2	40.90			67.61			50.97		

Hypothesis	**18 month**			**21 month**			**24 month**		
	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value	Log Likelihood	Trace test	p-value
rank=0	51.34	24.62	0.012	51.79	15.47	0.200	52.88	13.33	0.380
rank≤1	59.99	7.31	0.112	57.86	3.33	0.580	57.54	4.02	0.480
rank≤2	63.65			59.53			59.55		

Note: Cointegration rank determination with an intercept restricted in the cointegrating regression. Critical values are based on Johansen (1995, Table 2) and Doornik (1998) and the tests are performed at 5% level of significance. The Table summarises results from estimating a cointegrated VAR model for the logarithmic prices for the future spot and futures contracts across different maturities we estimate a bivariate VAR model for the sample period November 2015 to February 2016 at weekly frequency, fitting three lags and including an intercept in the cointegrating regression. The likelihood ratio test has the same asymptotic distribution as in Johansen (1995, Chap.6) even in the presence of an explosive root.

Table 11: Characteristic Roots

Futures Contracts	Characteristic Roots			
	July 2007-July 2008		November 2015-February 2016	
	Contemporaneous Series	Non-contemporaneous Series (Lags)	Contemporaneous Series	Non-contemporaneous Series (lags)
1 month	1.0107	0.9854	1.0350	0.9219
3 month	1.0162	0.9615+0.0364i	1.0570	1.2670
6 month	1.0226	1.0544	0.9837	0.9178+0.3787i
9 month	1.0260	1.0178	0.8352	1.0664
12 month	1.0274	1.0366	0.7715+0.1912i	0.8988+0.1465i
15 month	1.0283	1.0215	0.7982+0.2343i	0.8991
18 month	1.0288	1.0235	0.7767+0.2259i	0.8919+0.2411i
21 month	1.0290	1.0214	0.7479+0.2115i	1.0810
24 month	1.0289	1.0248	0.7149+0.1753i	0.6532+0.1326i

Note: The characteristic roots are the estimated roots of the characteristic polynomial for Equation (4.1) for the periods July 2007 to July 2008 and November 2015 to February 2016 using both contemporaneous and non-contemporaneous (lags) series. In the contemporaneous series, the WTI crude oil spot prices match with the futures contract prices on the same date whereas in the non-contemporaneous case, the WTI crude oil future contract prices match with the futures contract prices on the expiration date of the futures contracts.

Table 12: Granger Causality test

Futures Contracts	$H_{0,ss}$	$H_{0,sf}$	$H_{0,fs}$	$H_{0,ff}$
1 month	0.003*	0.000*	0.025*	0.000*
3 month	0.061	0.000*	0.077	0.000*
6 month	0.431	0.001*	0.908	0.006*
9 month	0.876	0.025*	0.637	0.234
12 month	0.683	0.188	0.334	0.612
15 month	0.381	0.596	0.081	0.661
18 month	0.243	0.969	0.048*	0.441
21 month	0.171	0.760	0.047*	0.323
24 month	0.138	0.628	0.028*	0.185

* statistically significant at 5%, $p - value < 0.05$

Note: Probabilities of rejection of the null hypothesis of no Granger causality. We consider a bivariate VAR(3) model with a constant and we use of the likelihood ratio test to investigate whether it is returns on spot prices causing returns on futures contract prices or vice versa. We fit three lags of spot and futures returns and we account for the case that the errors are heteroskedastic but uncorrelated by performing a heteroskedasticity robust likelihood ratio test. We test the hull hypothesis of no Granger causality across all possible restrictions. In particular, according to the null hypothesis $H_{0,ss}$, lag of returns on WTI crude oil spot prices do not Granger cause returns on WTI crude oil spot prices at time t. Respectively, according to the null hypothesis $H_{0,sf}$, lag of returns on WTI crude oil futures contract prices (of length n) do not Granger cause returns on WTI crude oil spot prices at time t whereas not rejecting the null hypothesis $H_{0,fs}$, would mean that lag of returns on WTI crude oil spot prices do not Granger cause returns on WTI crude oil futures contract prices (of length n) at time t. Finally, under to the null hypothesis $H_{0,ff}$, lag of returns on WTI crude oil futures contract prices (of length n) do not Granger cause returns on WTI crude oil futures contract prices (of length n) at time t.

Table 13: Unit Root Tests

Series/Test	*ADF*	*SADF*	*GSADF*
Panel A: test statistics			
$s_{t+1} - f_{t,1}$	-13.730	-3.031	-1.054
$s_{t+3} - f_{t,3}$	-6.622	-1.287	1.942
$s_{t+6} - f_{t,6}$	-4.371	-0.113	3.959*
$s_{t+9} - f_{t,9}$	-3.545	0.214	4.134*
$s_{t+12} - f_{t,12}$	-3.288	0.524	3.074*
$s_{t+15} - f_{t,15}$	-3.101	-0.104	3.108*
$s_{t+18} - f_{t,18}$	-2.761	0.254	2.958*
$s_{t+21} - f_{t,21}$	-2.648	0.283	2.484*
$s_{t+24} - f_{t,24}$	-2.503	0.242	2.387*
Panel B: critical values			
90%	-0.44	1.23	2.10
95%	-0.07	1.53	2.34
99%	0.60	2.03	2.79

* statistically significant at 5%

Note: Critical values of both *SADF* and *GSADF* tests are obtained from Monte Carlo simulations with $2,000$ replications (sample size 1,243 observations). We apply the ADF, SADF and GSADF tests to the difference between the future spot prices (s_{t+n}) and futures contract prices ($f_{t,n}$) where n is the contract length, across all different maturity contracts for the sample period September 1995 to July 2019.

Table 14: Bubble Explosion Date Stamping

Series	**BSADF**	**Series**	**BSADF**
s_t	2008M2-2008M7		
$f_{t,1}$	2008M2-2008M7	$s_{t+1} - f_{t,1}$	-
$f_{t,3}$	2008M2-2008M8	$s_{t+3} - f_{t,3}$	2008M10-2009M1
$f_{t,6}$	2007M10-2008M8	$s_{t+6} - f_{t,6}$	2008M10-2009M3
$f_{t,9}$	2007M10-2008M8	$s_{t+9} - f_{t,9}$	2008M11-2009M3
$f_{t,12}$	2007M10-2008M8	$s_{t+12} - f_{t,12}$	2008M11-2009M3
$f_{t,15}$	2007M10-2008M8	$s_{t+15} - f_{t,15}$	2008M12-2009M1
$f_{t,18}$	2007M10-2008M8	$s_{t+18} - f_{t,18}$	2008M12-2009M1
$f_{t,21}$	2004M4-2008M8	$s_{t+21} - f_{t,21}$	2008M12-2009M1
$f_{t,24}$	2004M4-2008M8	$s_{t+24} - f_{t,24}$	2008M12-2009M1

Note: Bubble date stamping application on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. We compute right-tailed finite sample critical values for the BSADF test using $2,000$ Monte Carlo replications. The test is performed at a 5% level of significance and a constant is included in the regression.

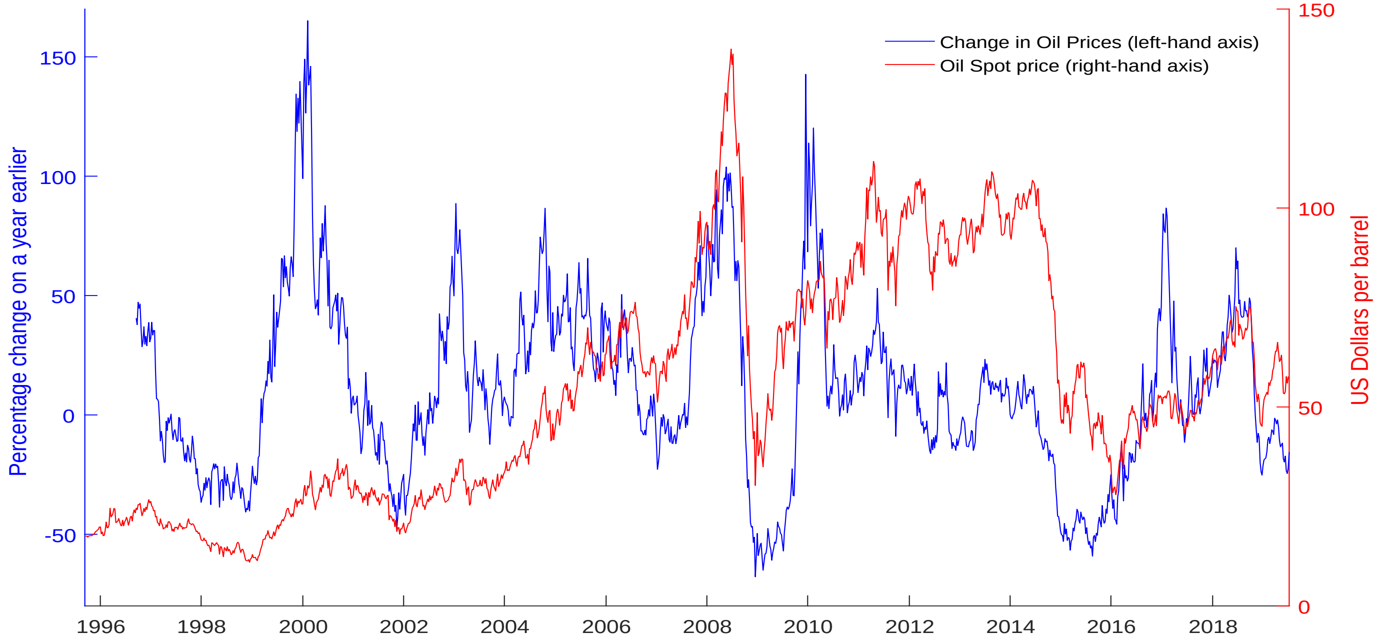
Table 15: Bubble Implosion Date Stamping

Series	**BSADF**	**Series**	**BSADF**
s_t	2016M3-2016M4		
$f_{t,1}$	2016M2-2016M5	$s_{t+1} - f_{t,1}$	-
$f_{t,3}$	2016M1-2016M5	$s_{t+3} - f_{t,3}$	-
$f_{t,6}$	2016M1-2016M5	$s_{t+6} - f_{t,6}$	-
$f_{t,9}$	2016M1-2016M4	$s_{t+9} - f_{t,9}$	2016M11-2016M12
$f_{t,12}$	2016M1-2016M4	$s_{t+12} - f_{t,12}$	-
$f_{t,15}$	2016M1-2016M3	$s_{t+15} - f_{t,15}$	2017M5-2017M7
$f_{t,18}$	2016M1-2016M3	$s_{t+18} - f_{t,18}$	2017M8-2017M10
$f_{t,21}$	2016M1-2016M3	$s_{t+21} - f_{t,21}$	2017M11-2017M12
$f_{t,24}$	2016M1-2016M3	$s_{t+24} - f_{t,24}$	2018M2-2018M3

Note: Bubble implosion date stamping application on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. We apply the BSADF test in the reverse order of the series that represents the difference between the future spot price and the futures contract price across futures contracts with different maturities so that if $y_t = s_{t+n} - f_{t,n}$ where n is the contract length then $y_t^* = y_{T+1-t}$ for $t = 1, 2, ..., T$. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. We compute right-tailed finite sample critical values for the BSADF test using $2,000$ Monte Carlo replications. The test is performed at a 5% level of significance and a constant is included in the regression.

4.9 Figures

Figure 22: WTI crude oil prices

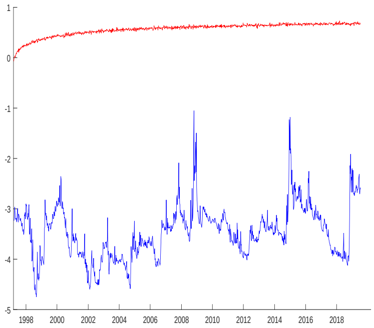


WTI crude oil spot prices together with the percentage change of WTI crude oil price on a year earlier.

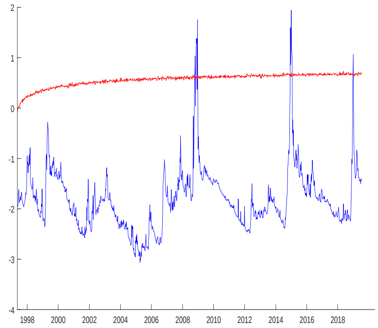
Figure 23: Johansen Cointegration Test on WTI crude oil prices

Note: Johansen cointegration analysis across different subsamples recursively. From June 2004, when the oil price run-up started, to July 2008 when the oil price collapsed, crude oil spot prices are cointegrated with all future contracts prices except the three month one. Additionally, crude oil spot prices are cointegrated between November 2015 and February 2016 marking the period of the 2014-2015 crude oil price collapse.

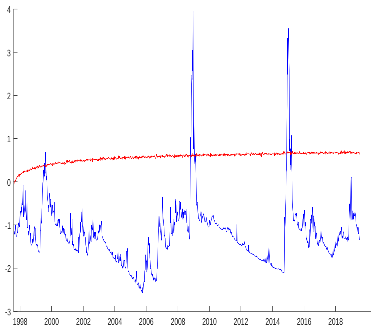
Figure 24: Bubble Explosion Date Stamping



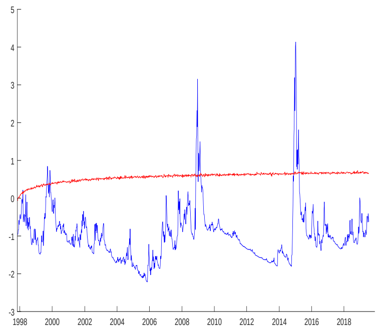
(a) $s_{t+1} - f_{t,1}$



(b) $s_{t+3} - f_{t,3}$



(c) $s_{t+6} - f_{t,6}$

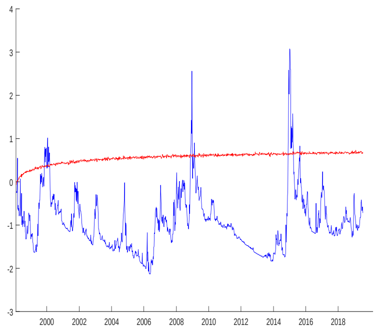


(d) $s_{t+9} - f_{t,9}$

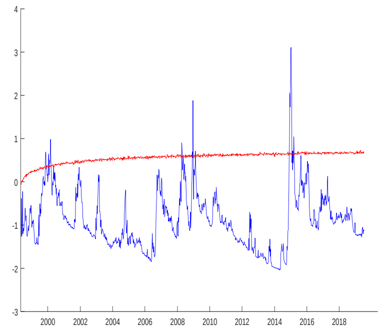
BSADF test: —— critical value sequence at 95%: ——

Note: The right-sided critical 95% values are approximated using Monte Carlo simulations with 2,000 replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Bubble date stamping is performed on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. The test is performed at a 5% level of significance and a constant is included in the regression.

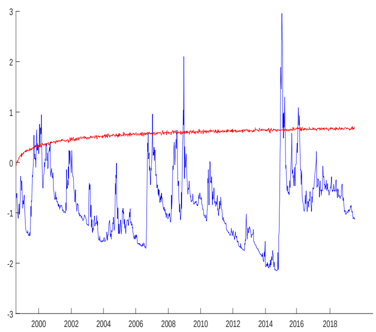
Figure 25: Bubble Explosion Date Stamping



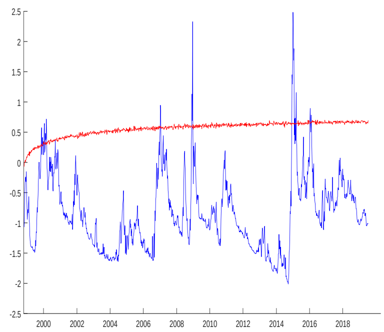
(a) $s_{t+12} - f_{t,12}$



(b) $s_{t+15} - f_{t,15}$



(c) $s_{t+18} - f_{t,18}$

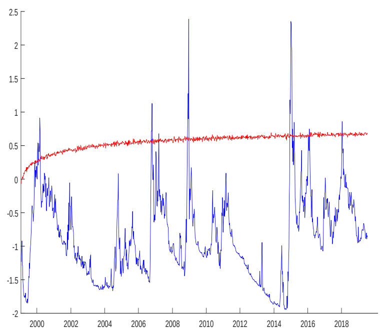


(d) $s_{t+21} - f_{t,21}$

BSADF test: ——— critical value sequence at 95%: ———

Note: The right-sided critical 95% values are approximated using Monte Carlo simulations with 2,000 replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Bubble date stamping is performed on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. The test is performed at a 5% level of significance and a constant is included in the regression.

Figure 26: Bubble Explosion Date Stamping

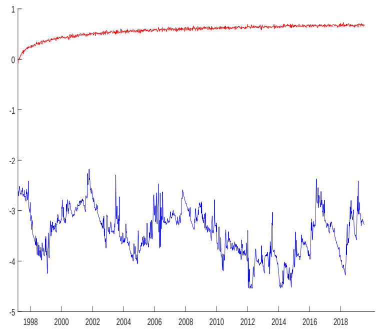


(a) $s_{t+24} - f_{t,24}$

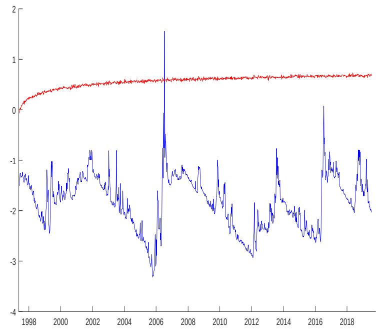
BSADF test: —— critical value sequence at 95%: ——

Note: The right-sided critical 95% values are approximated using Monte Carlo simulations with $2,000$ replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Bubble date stamping is performed on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. The test is performed at a 5% level of significance and a constant is included in the regression.

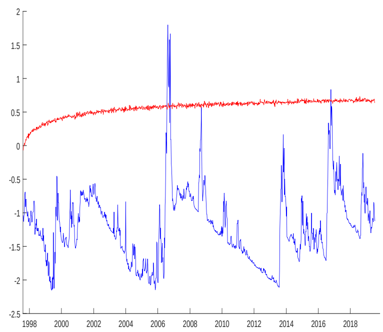
Figure 27: Bubble Implosion Date Stamping



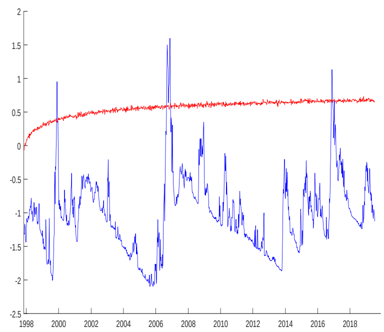
(a) $s_{t+1} - f_{t,1}$



(b) $s_{t+3} - f_{t,3}$



(c) $s_{t+6} - f_{t,6}$

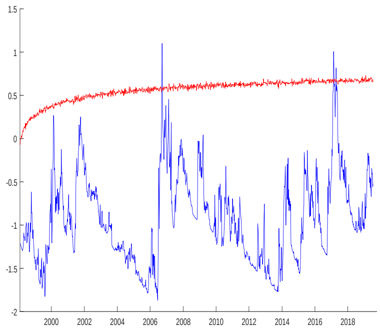


(d) $s_{t+9} - f_{t,9}$

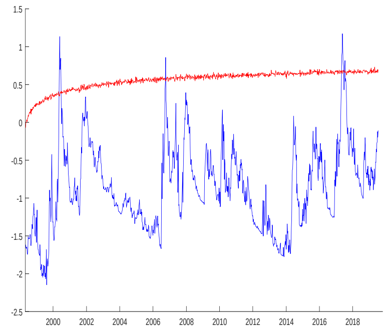
BSADF test: —— critical value sequence at 95%: ——

Note: The right-sided critical 95% values are approximated using Monte Carlo simulations with 2,000 replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Bubble date stamping is performed on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. The test is performed at a 5% level of significance and a constant is included in the regression.

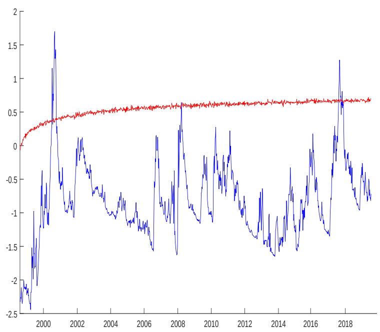
Figure 28: Bubble Implosion Date Stamping



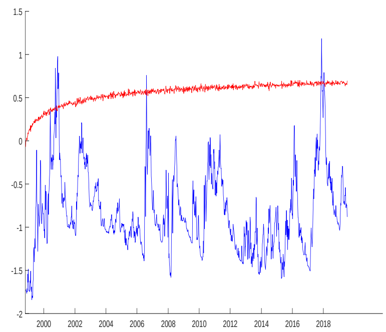
(a) $s_{t+12} - f_{t,12}$



(b) $s_{t+15} - f_{t,15}$



(c) $s_{t+18} - f_{t,18}$

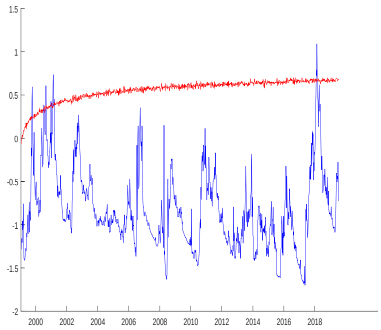


(d) $s_{t+21} - f_{t,21}$

BSADF test: —— critical value sequence at 95%: ——

Note: The right-sided critical 95% values are approximated using Monte Carlo simulations with 2,000 replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Bubble date stamping is performed on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. The test is performed at a 5% level of significance and a constant is included in the regression.

Figure 29: Bubble Implosion Date Stamping



(e) $s_{t+24} - f_{t,24}$

BSADF test: ——— critical value sequence at 95%: ———

Note: The right-sided critical 95% values are approximated using Monte Carlo simulations with 2,000 replications. The minimum window size is determined as in Equation (2.41) where T is the sample size of the observations as outlined by Phillips et al. (2015). Bubble date stamping is performed on the WTI crude oil spot and futures contract logarithmic prices over the period September 1995 to July 2019 at a weekly frequency, constituting 1243 observations. The futures contracts maturity ranges from one month to three, six, nine, twelve, fifteen, eighteen, twenty one and twenty four months and the futures contracts expire on the third business day prior to the twenty fifth calendar day of the month prior to the delivery month. The test is performed at a 5% level of significance and a constant is included in the regression.

5 Concluding Remarks

In this thesis we have proposed test procedures that result in early identification of bubble episodes in financial time series while dealing with the lower power of extant unit root tests. Additionally, we provide empirical evidence that in the presence of speculative bubbles, two series can be cointegrated so their linear combination does not contain a unit root while at the same time there is an explosive root in the system. This is in contrast to what is commonly mentioned in the literature that the existence of speculative bubbles in two series implies no cointegration between the two series.

In the second chapter of this thesis, a bootstrap unit root test that includes covariates has been proposed. It has been shown that the inclusion of relevant covariates in the conventional Augmented Dickey Fuller test regression leads to improved size control, while offering significant power gains, when an end-of-sample explosive episode is present. Concentrating on identifying explosive episodes that occur at the end of the sample, as the detection of ongoing bubbles is of most importance to practitioners, our proposed test has been applied in a recursive window framework as suggested by Phillips et al. (2015). Dealing with potential bias we have applied a bootstrap procedure of the proposed covariate test to ensure the asymptotic validity of the critical values drawn from the bootstrap distribution of the test. Simulation results have shown the ability to control size while offering great power gains in finite samples relative to extant tests. In particular, the CBSADF* test suffers less severe size distortions compared to conventional tests that do not utilise a bootstrap procedure or omit relevant covariates, whilst displaying significantly better power properties as well.

We have conducted empirical work that put emphasis on whether the proposed test can be effective as an early warning mechanism, indicating the possibility of a bubble episode to occur, contributing to structuring macroprudential policy. We have examined the effectiveness of our proposed test on earlier detection of historical bubbles in the S&P 500 price dividend series compared to the tests of Phillips et al. (2015), utilising the Moody's Seasoned Aaa and Baa Corporate Bond Yields, the Ten-Year Treasury Rate and the Volatility Index (VXO) as covariates. We have found that our proposed test results in the earlier detection of two major explosive episodes: Black Monday of October 1987 and the dot-com bubble.

In the third chapter of this thesis, a wild bootstrap procedure of the right-tailed Dickey-Fuller recursive unit root tests of Phillips et al. (2015) has been proposed to test for market efficiency in the commodity markets. This approach resembles the behaviour of heteroskedastic errors in a financial time series when there are structural breaks, regime changes or volatility breaks offering robust critical values. In the simulations it has been found that our proposed test offers better size control and power performance in finite samples. It has been shown that our proposed wild bootstrap test has less size distortions compared to the non-bootstrap test while the power performance has been significantly improved.

In the empirical exercise we have applied the proposed and extant tests on the difference between the WTI crude oil future price and the price of nine futures contracts across different maturities as in Pavlidis et al. (2018) over the period September 1995 to July 2019 at weekly and monthly frequency. Our proposed test has identified the 2007-2008 oil price run-up and the 2014-2015 oil price collapse while the Phillips et al. (2015) test has either not identified an episode at all, or identified the origination day of the episode with delay, reflecting the superior power of our proposed wild bootstrap test to effectively identify episodes of non-stationarity that occur at the end of the sample. The proposed test has suggested periods of market inefficiency prior to the existence of the bubble episode as identified by the conventional tests of Phillips et al. (2015).

In the forth chapter, we examined empirically whether two financial series can be cointegrated and yet, their linear combination contain an explosive component. As a result, the VAR approach developed by Johansen (1991) allows testing for cointegration while examining whether at least one of the variables has an explosive characteristic root. Extending the Johansen (1988) approach to allow for explosive roots in the cointegrating system as suggested by Nielsen (2010) can offer valuable information since it allows to test for cointegration across different series while simultaneously testing whether the series contain any explosive components, allowing to perform the cointegration analysis of Johansen (1991) even in the presence of explosive behaviour in the related series.

We have utilised Johansen's cointegration rank test to analyse an explosive episode in the WTI crude oil market between July 2007 and July 2008 as well as an oil price collapse between November 2015 and February 2016 at a weekly frequency. We have provided evidence that when we use contemporaneous crude oil prices for spot and futures contracts there is a single explosive root in the cointegrated VAR model between the crude oil spot prices and all futures contracts while at the same time the series are cointegrated, whereas when we match the actual future spot price with the futures contract price, there is a single explosive root in the cointegrated VAR model between the crude oil spot prices and the six month, twelve month and eighteen month futures contract while at the same time the series are cointegrated for that time period. Therefore, we have concluded that the series of the crude oil spot prices and futures contracts coexplode and their linear relationship is stationary for the period July 2007 to July 2008.

When we test for the oil price collapse between November 2015 and February 2016 using contemporaneous series, we find a single explosive root in the cointegrated VAR model between the crude oil spot prices and the one month futures contract and between the crude oil spot prices and the three month futures contract while at the same time the series are cointegrated during this period. When using non-contemporaneous series, we conclude that there is a single explosive root in the cointegrated VAR model between the crude oil spot prices and the three month futures contract while at the same time the two series are cointegrated for that time period. Therefore, the series of the crude oil spot prices and futures contracts coexplode and their linear relationship is stationary from November 2015 to December 2016.

It has been found that during the periods July 2007 to July 2008 and November 2015 to February 2016 both oil prices of spot and futures contracts are $I(1,x)$ processes and the two variables cointegrate and therefore their linear combination is an $I(0)$ process for these periods. As an extension of our study, following Pavlidis et al. (2017) we have applied a date-stamping procedure to the difference between the future spot prices and the futures contract prices that results in a delayed identification of the origination date of the bubble oil episode of 2007-2008 and the oil price collapse of 2014-2015 and suggests no statistical evidence of explosiveness from July 2007 to July 2008 and November 2015 to February 2016 (in the reverse series) in accordance to our findings that during that period (future) spot prices and futures contract prices are cointegrated and as their VAR model contains, for some futures contracts, characteristic roots we have concluded that oil prices of the spot and futures contracts coexplode.

5.1 Future Research

The proposed bootstrap unit root testing for explosive behaviour using covariates may constitute a conservative and strict tool of macro-prudential policy and surveillance. Macroprudential regulation may focus on dealing with bubble episodes using tools that are specifically structured to do so (such as countercyclical capital requirements, credit constraints, credit-to-GDP ratio monitoring and margin requirements, see Borio 2003) than monetary policy instruments that might fuel the bubble. Overall, there is a great challenge for both theorists and empirical researchers to understand the magnitude of asset price bubbles, investigate their origin and causes and track their development.

A potential avenue for future research with respect to the second chapter would be to consider a multi-covariate model, theoretically, to examine the size and power performance of our proposed covariate bootstrap BSADF test as including more information from related series could, potentially, contribute, empirically, to the earlier identification of bubble episodes in real time while offering even greater size control and power gains.

An immediate avenue for further research arising from the third chapter would be to allow for the possibility of non-stationary volatility in the innovations. To allow for non-stationary volatility we could consider volatility breaks in our proposed wild bootstrap BSADF test as in Harvey et al. (2017). We envisage that allowing the innovations of the series to exhibit non-stationary volatility could be considered, without affecting the good size performance of our proposed test significantly and could, possibly when applied for date-stamping processes, lead to a more precise estimation of the origination and termination dates of episodes of non-stationarity. We leave this for future research.

The wide swings in crude oil prices during the early 2000s have attracted the interest of professionals, regulators, academics and policy makers. Many argue that crude oil price fluctuations have mainly been attributed to reasons related to oil demand and supply shocks with the popular view that the remarkable price increase of 2007-2008 is mainly related to increase in demand. Although there is a growing consensus that speculative activity together with the financialisation of the oil futures market might have caused a speculative bubble that subsequently collapsed in mid-2008, identifying speculative bubbles in the oil market can be rather inconclusive since the fundamental price of oil cannot be observed. Therefore, any statistical evidence of explosive behaviour can either be attributed to a misspecified model for fundamentals or the existence of speculative bubbles (the joint hypothesis problem) or both (Gürkaynak 2008).

One might argue that examining variables in a bivariate framework might offer significant advantages as bubble episodes emerging in the futures market might be transmitted in the spot market causing speculative bubbles and thus cointegration and coexplosive analysis can be proven valuable in bubble identification. A natural avenue for future research would be to consider applying time-varying causality tests to examine the casual impact of crude oil spot prices to futures contract prices and vice versa utilising forward recursive and rolling window tests as in Shi et al. (2018). Additionally, market expectations could be, potentially, utilised in that framework as in Pavlidis et al. (2018) who use market expectations for WTI crude oil future contract prices as a fundamental price, although this data is proprietary. Moreover, it would be interesting to apply a recursive bootstrap algorithm to determine the cointegration rank in the coexplosive VAR model of Nielsen (2010) to ensure the bootstrap statistics converge weakly to the usual asymptotic distributions and the probability of choosing the rank smaller than the true one converges to zero as in Swensen (2006).

The significance of crude oil for the real economy has widely been acknowledged and the magnitude of the oil price spikes can be great. Speculative bubbles in the oil futures market might require stricter regulation of speculation to minimise the impact of oil price collapses on the real economy. Furthermore, there is a great need for deeper understanding of the mechanisms asset bubbles formation, by the central banks and policy makers, as well as information on how they grow, collapse and contaminate other markets and the real economy. Fiscal regulators and institutional surveillance mechanisms require tools with low false detection rate to implement macro-prudential policy implementation to address bubble episodes in financial markets (Phillips et al. 2015).

The question whether central banks should intervene when a speculative bubble grows or wait until the crash takes place or whether these bubbles have rational or behavioural determinants is still left to be answered. There is great scope for further research.

A Appendix The Limit Theory of the CGSADF and the bootstrap CGSADF test

Proofs

Proof of Theorem 2.1

(a) If we denote a $CADF$ unit root test statistic applied to the subsample of data from $t = \lfloor r_1 T \rfloor, ..., \lfloor r_2 T \rfloor$ as t_{r_1,r_2} then following Chang, Sickles and Song (2017) we first define

$$A_{r_1,r_2} = \sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}\varepsilon_t - \left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t'\right)\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t\varepsilon_t\right)$$

$$B_{r_1,r_2} = \sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}^2 - \left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t'\right)\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t y_{t-1}\right)$$

$$C_{r_1,r_2} = \sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} \varepsilon_t^2 - \left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} \varepsilon_t z_t'\right)\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t\varepsilon_t\right)$$

If we define $r = r_2 - r_1$, then from Lemma 2.1 of Park and Phillips (1989) we know that

$$\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t' = O_p(r), \quad \sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t\varepsilon_t = O_p(r^{1/2}) \text{ and } \sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t' = O_p(r).$$

therefore

$$\left|\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t'\right)\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t\varepsilon_t\right)\right| \leq \left|\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t'\right|\left|\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\right|\left|\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t\varepsilon_t\right| = O_p(r^{1/2}),$$

$$\left|\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t'\right)\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t y_{t-1}\right)\right| \leq \left|\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} y_{t-1}z_t'\right|\left|\left(\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t z_t'\right)^{-1}\right|\left|\sum_{t=\lfloor r_1 T \rfloor}^{\lfloor r_2 T \rfloor} z_t y_{t-1}\right| = O_p(r),$$

$$\left|\left(\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor}\varepsilon_t z_t'\right)\left(\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} z_t z_t'\right)^{-1}\left(\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} z_t\varepsilon_t\right)\right| \leq \left|\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor}\varepsilon_t z_t'\right|\left|\left(\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} z_t z_t'\right)^{-1}\right|\left|\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} z_t\varepsilon_t\right| = o_p(r).$$

Thus,

$$(\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}A_{r_1,r_2} = (\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y_{t-1}\varepsilon_t + o_p(1),$$

$$(\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-2}B_{r_1,r_2} = (\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-2}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y_{t-1}^2 + o_p(1),$$

$$(\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}C_{r_1,r_2} = (\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} \varepsilon_t^2 + o_p(1).$$

Furthermore,

$$A_{r_1,r_2}B_{r_1,r_2}^{-1/2} \xrightarrow{d} \sigma_\varepsilon \frac{\int_{r_1}^{r_2} Q(s)dP(s)}{\left(\int_{r_1}^{r_2} Q(s)^2 ds\right)^{1/2}}$$

and therefore it follows from Park and Phillips (1989) that

$$\hat{\sigma}_{r_1,r_2}^2 = E(\varepsilon_t^2) + O_p(r^{-1}) = (\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor}(\varepsilon_t - \bar{\varepsilon}_{r_1,r_2})^2 + O_p(r^{-1}) = \sigma_\varepsilon^2 + o_p(1).$$

Under the null hypothesis of $\delta = 0$ we thus have that

$$CGSADF = \sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} CADF_{r_1}^{r_2} = \sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} \frac{1}{\hat{\sigma}_{r_1,r_2}}\frac{A_{r_1,r_2}}{B_{r_1,r_2}^{1/2}} =$$

$$\sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} \frac{1}{\hat{\sigma}_{r_1,r_2}}\left(\frac{(\lfloor r_2T\rfloor-\lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y_{t-1}\varepsilon_t}{\left((\lfloor r_2T\rfloor-\lfloor r_1T\rfloor)^{-2}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y_{t-1}^2\right)^{1/2}}\right) + o_p(1) \xrightarrow{d} \sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} \frac{\int_{r_1}^{r_2} Q(s)dP(s)}{\left(\int_{r_1}^{r_2} Q(s)^2 ds\right)^{1/2}}$$

The proof of Theorem 2.1 (b)-(c) is obtained in a similar manner to that of Theorem 2.1 (a).

Proof of Theorem 2.2

(a) The stochastic orders for the bootstrap sample moments appearing in the definitions of the bootstrap $CGSADF^*$ test are easily obtained. Following Chang, Sickles and Song (2017) we first define

$$A^*_{r_1,r_2} = \sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}\varepsilon^*_t - \left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}{z^*}'_t\right)\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t\right)$$

$$B^*_{r_1,r_2} = \sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^{*2}_{t-1} - \left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}z^{*\prime}_t\right)\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t y^*_{t-1}\right)$$

$$C^*_{r_1,r_2} = \sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} \varepsilon^{*2}_t - \left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} \varepsilon^*_t z^{*\prime}_t\right)\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t\right)$$

If we define $r = r_2 - r_1$ also we have that

$$\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t = O^*_p(r),\quad \sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t = O^*_p(r^{1/2}) \text{ and } \sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}z^{*\prime}_t = O^*_p(r).$$

therefore

$$\left|\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}z^{*\prime}_t\right)\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t\right)\right| \le \left|\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}z^{*\prime}_t\right|\left|\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\right|\left|\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t\right| = O^*_p(r^{1/2}),$$

$$\left|\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}z^{*\prime}_t\right)\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t y^*_{t-1}\right)\right| \le \left|\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y_{t-1}z^{*\prime}_t\right|\left|\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\right|\left|\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t y_{t-1}\right| = O^*_p(r),$$

$$\left|\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} \varepsilon^*_t z^{*\prime}_t\right)\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t\right)\right| \le \left|\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} \varepsilon^*_t z^{*\prime}_t\right|\left|\left(\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t z^{*\prime}_t\right)^{-1}\right|\left|\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} z^*_t\varepsilon^*_t\right| = o^*_p(r).$$

Consequently,

$$(\lfloor r_2 T\rfloor - \lfloor r_1 T\rfloor)^{-1}A^*_{r_1,r_2} = (\lfloor r_2 T\rfloor - \lfloor r_1 T\rfloor)^{-1}\sum_{t=\lfloor r_1 T\rfloor}^{\lfloor r_2 T\rfloor} y^*_{t-1}\varepsilon^*_t + o^*_p(1),$$

$$(\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-2}B^*_{r_1,r_2} = (\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-2}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y^{*2}_{t-1} + o^*_p(1),$$

$$(\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}C^*_{r_1,r_2} = (\lfloor r_2T\rfloor - \lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} \varepsilon^{*2}_{t} + o^*_p(1).$$

Furthermore,

$$A^*_{r_1,r_2}B^{*-1/2}_{r_1,r_2} \xrightarrow{d^*} \sigma_{\varepsilon^*}\frac{\int_{r_1}^{r_2} Q(s)dP(s)}{\left(\int_{r_1}^{r_2} Q(s)^2 ds\right)^{1/2}}$$

and therefore if we define $\hat{\sigma}^{*2}_{r_1,r_2}$ to be the bootstrap counterpart of $\hat{\sigma}^{2}_{r_1,r_2}$, then it follows from Park and Phillips (1989) that

$$\hat{\sigma}^{*2}_{r_1,r_2} = E(\varepsilon^{*2}_t)+O^*_p(r^{-1}) = (\lfloor r_2T\rfloor-\lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor}(\varepsilon^*_t-\bar{\varepsilon}^*_{r_1,r_2})^2+O^*_p(r^{-1}) = \sigma^2_{\varepsilon^*}+o^*_p(1).$$

Under the null hypothesis of $\delta^* = 0$, the stated limit distribution of $CGSADF^*$ follows from Equation (2.15) as

$$CGSADF^* = \sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} CADF^{*r_2}_{r_1} = \sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} \frac{1}{\hat{\sigma^*}_{r_1,r_2}}\frac{A^*_{r_1,r_2}}{B^{*1/2}_{r_1,r_2}} =$$

$$\sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} \frac{1}{\hat{\sigma^*}_{r_1,r_2}}\left(\frac{(\lfloor r_2T\rfloor-\lfloor r_1T\rfloor)^{-1}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y^*_{t-1}\varepsilon_t}{\left((\lfloor r_2T\rfloor-\lfloor r_1T\rfloor)^{-2}\sum_{t=\lfloor r_1T\rfloor}^{\lfloor r_2T\rfloor} y^{*2}_{t-1}\right)^{1/2}}\right) + o*_p(1) \xrightarrow{d} \sup_{\substack{r_2\in[r_0,1]\\ r_1\in[0,r_2-r_0]}} \frac{\int_{r_1}^{r_2} Q(s)dP(s)}{\left(\int_{r_1}^{r_2} Q(s)^2ds\right)^{1/2}}$$

The proof of Theorem 2.2 (b)-(c) is obtained in a similar manner to that of Theorem 2.2 (a).

References

Abadir, K. M., & Lucas, A. (2000). Quantiles for t-statistics based on M-estimators of unit roots. Economics Letters, 67(2), 131-137.

Abreu, D., & Brunnermeier, M. K. (2003). Bubbles and crashes. Econometrica, 71.1 , 173-204.

Akerlof, G. A., & Shiller, R. J. (2009). Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism. Princeton: Princeton University Press.

Andrews, D.W.K. (2003). End-of-sample instability tests. Econometrica 71, 1661-1694.

Andrews, D.W.K. and Kim, J.Y. (2006). Tests for cointegration breakdown over a short time period. Journal of Business and Economic Statistics 24, 379-394.

Aristidou, C., Harvey, D., & Leybourne, S. J. (2017). The impact of the initial condition on covariate augmented unit root tests. Journal of Time Series Econometrics, 9(1).

Astill, S., Harvey, D. I., Leybourne, S. J., & Taylor, A. R. (2017). Tests for an end-of-sample bubble in financial time series. Econometric Reviews, 36:6-9.

Avery, C., & Zemsky, P. (1998). Multidimensional uncertainty and herd behavior in financial markets. American Economic Review, 724-748.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. Econometrica, 47-78.

Baillie, R. T. (1989). Econometric tests of rationality and market efficiency. Econometric Reviews 8, 151-186.

Bettendorf, T., & Chen, W. (2013). Are these bubbles in the Sterling-dollar exchange rate? New evidence from sequential ADF tests. Economic Letters, 120(2), 350-353.

Bharghava, A. (1986). On the theory of testing for unit roots in observed time series. Review of economic Studies, 53, 369-384.

Blanchard, O. J. (1979). Speculative bubbles, crashes and rational expectations. Economic letters, 3(4), 387-389.

Blanchard, O. J., & Watson, M. W. (1982). Bubbles, rational expectations and financial markets., in: P. Wachtel, ed., Crises in the economic and financial structure, Lexington, MA.

Bohl, M. T., & Sicklos, P. L. (2004). The present value model of US stock prices redux: a new testing strategy and some evidence. The Quarterly Review of Economics and Finance, 44(2), 208-223.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. Journal of econometrics, 73(1), 151-184.

Borio, C. (2003). Towards a macroprudential framework for financial supervision and regulation? CESifo Economic Studies, 49(2), 181-215.

Brockwell, P. J., & Davis, R. A. (1991). Time Series: Theory and Methods. New York: Springer-Verlag.

Busetti, F., & Taylor, A. R. (2004). Tests of Stationarity Against a Change in Persistence. Journal of Econometrics, 123, 33-66.

Calverley, J. (2004). Bubbles and how to survice them. Nicholas Brealey Publishing.

Campbell, J. Y., Lo, A. W. C., & MacKinlay, A. C. (1997). The Econometrics of financial markets. Princeton University Press.

Campbell, J. Y., & Perron, P. (1991). Pitfalls and opportunities: what macroeconomists should know about unit roots. NBER macroeconomics annual, 6, 141-201.

Campbell, J. Y., & Shiller, R. (1987). Cointegration and tests of present value models. Journal of Political Economy, 95, 1062-1088.

Campbell, J. Y., & Shiller, R. (1988). Stock prices, earnings, and expected dividends. Journal of Finance, 43, 661-676.

Campbell, J. Y., & Shiller, R. (1989). The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies, 1(3), 195-228.

Caporale, G. M., & Pittis, N. (1999). Unit root testing using covariates: some theory and evidence. Oxford Bulletin of Economics and Statistics, 61(4), 583-595.

Cavaliere, G., & Taylor, A. R. (2008). Bootstrap unit root tests for time series with nonstationary volatility. Econometric Theory, 24, 43-71.

Cavaliere, G., & Taylor, A. R. (2009). Heteroskedastic Time Series with a Unit Root. Econometric Theory, 25(5), 1228-1276.

CFTC. (2008). Staff Report on Commodity Swap Dealers & Index Traders with Commission Recommendations. Commodity Futures Trading Commission (September).

Chang, Y., Sickles, R. C., & Song, W. (2017). Bootstrapping unit root tests with covariates. Econometric Reviews, 36:1-3, 136-155.

Christiano, L., Ilut, C. L., Motto, R., & Rostagno, M. (2008). Monetary Policy and Stock Market Boom-Bust Cycles. European Central Bank Working Paper No 955.

Charemza, W. W., & Deadman, D. F. (1995). Speculative bubbles with stochastic explosive roots: the failure of unit root testing. Journal of Empirical Finance, 2(2), 153-163.

Craine, R. (1993). Rational bubbles: A test. Journal of Economic Dynamics and Control, 17, 829-846.

Cuñado, J., Gil-Alama, L. A., & Perez de Gracia, F. (2005). A test for rational bubbles in the NASDAQ stock index: A fractionally integrated approach. Journal of Banking and Finance, 29, 2633-2654.

Demetrescu, M., Kuzin, V., & Hassler, U. (2008). Long Memory testing in the Time Domain. Econometric Theory, 24, 176-215.

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

Domowitz, I., & El-Gamal, M. (2001). A consistent nonparametric test of ergodicity for time series with applications. Journal of Econometrics, 102, 365-398.

Doornik, J. A. (1998). Approximations to the asymptotic distribution of cointegration tests. Journal of Economic Surveys 12, 573–93.

Elliott, G. (1998). On the robustness of cointegration methods when regressors almost have unit roots. Econometrica, 66, 149-158.

Elliott, G., and Jansson, M. (2003). Testing for unit roots with stationary covariates. Journal of Econometrics, 115, 75-89.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root testing. Econometric Theory, 813-836.

Engel, C., Mark, N. C., West, K. D., Rogoff, K., & Rossi, B. (2007). Exchange rate models are not as bad as you think [with comments and discussion]. NBER macroeconomics annual, 22, 381-473.

Engel, C., & West, K. D. (2005). Exchange rates and fundamentals. Journal of political Economy, 113(3), 485-517.

Engsted, T. (2006) Explosive bubbles in the cointegrated VAR model. Finance Research Letters 3, 154-162.

Engsted, T., & Nielsen, B. (2012). Testing for rational bubbles in a coexplosive vector autoregression. The Econometrics Journal, 15(2), 226-254.

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

Fantazzini, D. (2011). Fractionally integrated models for volatility: A review. In Nonlinear Financial Econometrics: Markov Switching Models, Persistence and Nonlinear Cointegration (pp. 104-123). Palgrave Macmillan, London.

Federal Reserve. 2016. Selected Interest Rates - H.15 (519) Release. Retrieved from https://www.federalreserve.gov/releases/h15/current/h15.pdf

Fleming, J., Ostdiek, B., & Whaley, R. E. (1995). Predicting stock market volatility: A new measure. Journal of Futures Markets, 15(3), 265-302.

Flood, R. P., & Garber, P. (1980). Market fundamentals versus price level bubbles: The first tests. Journal of Political Economy, 88, 745-770.

Flood, R., & Hodrick, R. (1986). Asset price volatility, bubbles and process switching. Journal of Finance, 41, 831-842.

Flood, R., Hodrick, R., & Kaplan, P. (1994). An evaluation of recent evidence on stock price bubbles. In R. Flood, and P. Garber, Speculative Bubbles, Speculative Attacks, and Policy Switching (pp. 105-133). Cambringe, MA: MIT Press.

Frömmel, M., & Kruse, R. (2012). Testing for a rational bubble under long memory. Quantitative Finance, 12(11), 1723-1732.

Froot, K., & Obstfeld, M. (1991). Intrinsic bubbles: the case of stock prices. American Economic Review, 81, 1189-1214.

Galbraith, J. K. (1997). The Great Crash (1929). Boston: Houghton Miffin Company.

Gilbert, C. L. (2010). Speculative influences on commodity futures prices 2006-2008. Discussion Paper 197, United Nations Conference on Trade and Development.

Greenspan, A. (1996). Minutes of the Federal Open Market Committee. Retrieved from: www.federal.reserve.gov/transcript/1996/19960924meeting.pdf

Gürkaynak, R. S. (2008). Econometric Tests of Asset Price Bubbles: Taking Stock. Journal of Economic Surveys, 22(1), 166-186.

Hall, S., Psaradakis, Z., & Sola, M. (1999). Detecting Periodically Collapsing Bubbles: A Markov-Switching Unit Root Test. Journal of Applied Econometrics., 14, 143-154.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57, 357-384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. Journal of Econometrics, 45, 39-70.

Hansen, E. B. (1995). Rethinking the Univariate Approach to Unit Root Testing; Using Covariates to Increase Power. Econometric Theory, 11, 1148-1171.

Harvey D. I., Leybourne S.J., & Sollis R, (2012). Testing for an asset price bubble and dating the exploding/collapsing regimes, working paper.

Harvey, I. D., Leybourne, S. J., & Sollis, R. (2015). Recursive right-tailed unit root tests for an explosive asset price bubble. Journal of Financial Econometrics , 13, 166-187.

Harvey, I. D., Leybourne, S. J., Sollis, R., & Taylor, R. A. (2017). Tests for explosive financial bubbles in the presence of non-stationary volatility. Journal of Empirical Finance, 40, 121-138.

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

Huang, W., Zheng, H., & Chia, W. M. (2010). Financial crises and interacting heterogeneous agents. Journal of Economic Dynamics and Control, 34(6), 1105-1122.

Hylleberg, S., Engle, R. F., Granger, C. W., & Yoo, B. S. (1990). Seasonal integration and cointegration. Journal of econometrics, 44(1), 215-238.

Johansen, S. (1988). Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 12, 231-254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551-1580.

Johansen, S. (1995). Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford: Oxford University Press.

Johansen, S., & Swensen, A. R. (1999). Testing exact rational expectations in cointegrated vector autoregressive models. Journal of Econometrics, 93(1), 73-91.

Johansen, S., & Swensen, A. R. (2004). More on testing exact rational expectations in cointegrated vector autoregressive models: Restricted constant and linear term. The Econometrics Journal, 7(2), 389-397.

Johansen, S., & Swensen, A. R. (2011). On a graphical technique for evaluating some rational expectations models. Journal of Time Series Econometrics, 3(1).

Kejriwal, M., Perron, P., & Zhou, J. (2013). Wald tests for detecting multiple structural changes in persistence. Econometric Theory, 29(2), 289-323.

Kindleberger, C. P. (1987). International capital movements. CUP Archive. Chicago.

Kindleberger, C. P., & Aliber, R. Z. (2005). Manias, panics, and crashes: a history of financial crises. 5^{th} ed. Hoboken NJ: John Wiley & Sons, Inc.

Koustas, Z., & Serletis, A. (2005). Rational bubbles or persistent deviations from market fundamentals? Journal of Banking and Finance, 29, 2523-2539.

Lee, J. H., & Phillips, P. C. (2015). Asset pricing with financial bubble risk. Journal of Empirical Finance, 38, 590-622.

Lei, V., Noussair, N. C., & Plott, C. (2001). Nonspeculative bubbles in experimental asset markets: Lack of common knowledge or rationality vs. actual irrationality. Econometrica, 69(4), 831-859.

LeRoy, S. F., & Porter, R. D. (1981). The present value relation: Tests based on implied variance bounds. Econometrica, 49, 555-574.

Leybourne, S. J. (1995). Testing for unit roots using forward and reverse Dickey-Fuller regressions. Oxford Bulletin of Economics and Statistics, 57(4), 559-571.

Leybourne, S. J., & Taylor, A. R. (2003). Seasonal unit root tests based on forward and reverse estimation. Journal of Time Series Analysis, 24(4), 441-460.

Lucas, R. E. (1978). Asset prices in an exchange economy. Econometrica, 46, 1429-1445.

Magdalinos, T., & Phillips, P. (2009). Limit Theory for Cointegrated systems with Moderately Integrated and Moderately explosive Regressors. Econometric Theory, 25, 482-526.

Mandelbrot, B. B. (1969). Long range linearity, locally Gaussian process, H-spectra and infinite variance. International Economic Review, 10, 82-111.

Nelson, C. R., & Plosser, C. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. Journal of Monetary Economics, 10, 139-162.

Nielsen, B. (2010). Analysis of coexplosive processes. Econometric Theory 26, 882-915.

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

Park, J. Y., & Phillips, P. C. (1989). Statistical inference in regressions with integrated processes: Part 2. Econometric Theory, 5(1), 95-131.

Parke, W. R. (1999). What is fractional integration?. Review of Economics and Statistics, 81(4), 632-638.

Pavlidis, E. G., Paya, I., & Peel, D. A. (2017). Testing for speculative bubbles using spot and forward prices. International Economic Review, 58(4), 1191-1226.

Pavlidis, E. G., Paya, I., & Peel, D. A. (2018). Using market expectations to test for speculative bubbles in the crude oil market. Journal of Money, Credit and Banking, 50(5), 833-856.

Phelps, E., & Zoega, G. (2001). Structural booms. Economic Policy, 16(32), 84-126.

Phillips, P. C. (2016). Modelling Speculative Bubbles with Diverse Investor Expectations. Research in Economics.

Phillips, P. C., & Magdalinos, T. (2008). Limit theory for explosive cointegrated systems. Econometric Theory, 24, 865-887.

Phillips, P. C., & Magdalinos, T. (2009). Unit root and cointegrating limit theory when initialization is in the infinite past. Econometric Theory, 25(6), 1682-1715.

Phillips, P. C., & Shi, S. P. (2018). Financial Bubble Implosion and Reverse Regression. Econometric Theory. 1-49.

Phillips, P. C., Shi, S., & Yu, J. (2015). Testing for Multiple Bubbles : Historical Episodes of Exuberance and Collapse in the SP500. International Economic Review, 56(4), 1043-1078.

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

Rosser, J. B. (2000). From Catastrophe to Chaos: A General Theory of Economic Discontinuities: Mathematics. Microeconomics and Finance, Vol. 1, Springer Science & Business Media.

Saporta, V., Tudela, M., & Trott, M. (2009). What can be said about the rise and fall in oil prices?, Bank of England Quarterly Bulletin 2009 Q3.

Sarno, L., & Taylor, M. P. (2003). The economics of exchange rates. Cambridge University Press.

Schotman, P. C., & van Dijk, H. K. (1991). On Bayesian routes to unit roots. Journal of Applied Econometrics, 6, 387-401.

Shi, S., Phillips, P. C., & Hurn, S. (2018). Change detection and the causal impact of the yield curve. Journal of Time Series Analysis, 39(6), 966-987.

Shiller, R. J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? American Economic Review, 71 , 421-436.

Shiller, R. J. (2015). Irrational Exuberance. 3^{rd} ed. New Jersey: Princeton University Press.

Shleifer, A. (2000). Inefficient Markets: An introduction to behavioural finance. Oxford : OUP Oxford.

Sibbertsen, P., & Kruse, R. (2009). Testing for a break in persistence under long-range dependencies. Journal of Time-Series Analysis, 30(3), 263-285.

Sornette, D. (2003b). Why Stock Markets Crash; Critical Events in Complex Financial Systems. New Jersey: Princeton University Press.

Sornette, D., & Cauwels, P. (2012). The illusion of the perpetual money machine. Notestein Academy White Paper Series.

Stock, J. H. (1994). Unit roots, structural breaks and trends. Handbook of econometrics, 4, 2739-2841.

Swensen, A. R. (2006). Bootstrap Algorithms for Testing and Determining the Cointegration Rank in VAR Models 1. Econometrica, 74(6), 1699-1714.

Taylor, J. (2007), Housing and Monetary Policy, No 13682, NBER Working Papers, National Bureau of Economic Research, Inc, https://EconPapers.repec.org/RePEc:nbr:nberwo:13682.

Tsvetanov, D., Coakley, J., & Kellard, N. (2016). Bubbling over! The behaviour of oil futures along the yield curve. Journal of Empirical Finance, 38, 516-533.

Van Norden, S. (1996). Regime switching as a test for exchange rate bubbles. Journal of Applied Econometrics, 11, 219-251.

Van Norden, S., and Vigfusson, R. (1998). Avoiding the pitfalls: can regime-switching tests reliably detect bubbles? Studies in Nonlinear Dynamics and Econometrics, 3, 1-22.

Vogel, H. L. (2009). Financial market bubbles and crashes. Cambridge University Press.

West, K. (1987). A Specification Test for Speculative Bubbles. The Quarterly Journal of Economics, 102, 553-580.

Westerlund, J. (2013). A computationally convenient unit root test with covariates, conditional heteroskedasticity and efficient detrending. Journal of Time Series Analysis, 34(4), 477-495.