University of Macedonia

MASTER THESIS

Revisiting cointegration tests: power versus frequency - some further Monte Carlo results

Author:
Georgios Papapanagiotou

Supervisor:
Theodore Panagiotidis

March 4, 2019

A thesis submitted in partial fulfillment of the requirements for the degree of
MASTER IN ECONOMICS

Interdepartmental Programme of Postgraduate Studies
in Economics

with specialization
in Applied Economics and Finance

Abstract

We study the effect of increasing the frequency of observations and the data span on alternative cointegration tests (Engle-Granger, Phillips-Ouliaris and Johansen), we consider systems with two, three and four variables via Monte Carlo simulations. We find that when both the data length and the frequency vary, the power of the tests depends more on the sample length. In addition, we explore the behaviour of unit root tests and the Engle-Granger and Johansen methods when explosive processes are included. The results show that the performance of the tests depends on the kind of the type of the explosion.

Acknowledgements

First and foremost I would like to thank my supervisor Prof. Theodore Panagiotidis for the helpful advice and suggestions. His continuous support and guidance helped the most to the improvement of this thesis. I would also like to thank my brother and my parents for their encouragement and moral support. Last but not least, I would like to thank Melina for her unconditional love and support.

Contents

List of Tables

Chapter 1

Introduction

In probability theory and statistics, a unit root is a feature of some stochastic processes (such as random walks) that can cause problems in statistical inference involving time series models. A linear stochastic process has a unit root if a root of the process's characteristic equation has an absolute value equal to 1. If the other roots of the characteristic equation lie inside the unit circle, that is, have a modulus (absolute value) less than one, then the first difference of the process will be stationary; otherwise, the process will need to be differenced multiple times to become stationary. Due to this characteristic, unit root processes are also called difference stationary. A process that has to be differenced k times to become stationary is also called integrated of order k, $(I(1))$.

In an influential paper, Nelson and Plosser (1982) provided statistical evidence that many US macroeconomic time series (like GNP, wages, employment, etc.) have stochastic trends. They also showed that unit root processes have non-standard statistical properties, so that conventional econometric theory methods do not apply to them. In fact, they showed that applying usual linear regressions non-stationary variables leads to the so called spurious regression.

Imagine the regression $y_t = a_0 + a_1 x_t + u_t$ and assume that x_t is an random walk and y_t is an independent random walk. Then the true value of a_1 is of course 0, but the limiting distribution of $\hat{a_1}$ is such that it converges to a function of Brownian motions. This is an example of a spurious regression. A very high value of the coefficient of determination (R^2) and a very low Durbin-Watson value indicate that one possibly is facing a regression where unit roots should be taken into account.

The phenomenon that non-stationary processes can have linear combinations that are stationary was called cointegration by Granger (1983), who used it for modelling long-run economic relations. Cointegration is a statistical property of a collection $(X_1, X_2, ..., X_k)$ of time series variables. First, all of the series must be integrated of order d. Next, if a linear combination of this collection is integrated of order zero, then the collection is said to be co-integrated. Formally, if (X, Y, Z) are each integrated of order d, and there exist coefficients a, b, c such that $aX + bY + cZ$ is integrated of order 0, then X, Y, and Z are cointegrated. Cointegration has become an important property in contemporary time series analysis. The paper by Engle and Granger (1987), which showed the equivalence of the error correction formulation and the phenomenon of cointegration, started a rapid development of the statistical and probabilistic analysis of the ideas.

There are three popular methods for testing for cointegration. The Engle-

Granger two-step method, the Phillips-Ouliaris test and the Johansen test. Each one has different properties, advantages and disadvantages. Since their development these methods have been well studied and are now an important tool in time series analysis.

The term unit root process is sometimes inaccurately used to describe explosive processes. A process is called explosive if its characteristic equation contains a root with a modulus greater than one. Although an explosive process looks similar to non-stationary one, it requires a more careful approach and analysis. Imagine for example an autoregressive one ((AR(1)) explosive process $x_t = \rho x_{t-1} + u_t$ where $\rho > 1$. One can easily understand how a shock in the disturbance term will affect the behaviour of such a process. Until recently it was considered unlikely explosive processes to arise in economics. However studies have shown that a financial crises is often preceded by an asset market growth or rampant credit growth.

The development of econometric tests that can detect explosive behaviour have attracted many researchers. A study by Gürkaynak (2005) showed that until that moment econometric detection of asset price bubbles cannot be achieved with a satisfying degree of certainty. Phillips et al. (2011) proposed a method for detecting a single bubble in asset price series. Although this method has been proven effective enough, it is insufficient when applied on time series which contain multiple bubbles. Phillips et al. (2013) provided a framework for testing and dating multiple bubble incidents in a series by extending the mechanisms of the Phillips et al. (2011).

The violation of usual asymptotic properties of cointegration tests caused by explosive behaviour lead the researchers to employ the $I(2)$ model for cointegration analysis when there were explosive roots in the data. This changed when Bent (2000) and Nielsen (2010) showed that the asymptotic results for the Johansen's test are valid for explosive growing variables.

In this study we examine the effects of temporal disaggregation in cointegration analysis. Our study consists of two main parts. In the first part we perform a large number of Monte Carlo simulation experiments to examine the effects of changing the frequency of observations and the total data span on the Engle-Granger, the Phillips-Ouliaris and the Johansen cointegration tests. Based mostly on Otero and Smith (2000) we examine the importance of the two properties in systems with three and four variables. We focus our study mostly on the Johansen test since it allows for more than one cointegrating relationship and it is the most used methodology when testing for cointegration in multivariate systems. The extant literature is rich in studies that investigate the effects of temporal disaggregation. However, all papers only consider two variables. We contribute to the existing literature by considering cases of three and four variables.

In the second part we focus our research on explosive time series. We examine the behaviour of the SADF and the GSADF tests when both the data span and the frequency vary. We consider three different explosive models proposed by Phillips et al. (2013), Blanchard (1979) and Evans (1991). We then employ the Johansen and the Engle-Granger test in systems which exhibit explosive behaviour.

The main results of our study can be summarized along these lines. Considering the three cointegration tests, the ability of the three examined methods to detect cointegration is based more on the total sample length than on the frequency of observations when all variables are integrated of order one. When explosive processes enter the system we observe the following pattern. In the presence of a sole bubble, both methods (the Phillips-Ouliaris is not considered in this part of the study) perform better when a longer period is used. However when they deal with periodically collapsing bubbles the power of both tests depends more on the frequency of observations. Finally, the SADF and GSADF tests produce better results when applied on high frequency samples rather than a longer data length.

The analysis was conducted using the **R** statistical language. **R** is a programming language and free software environment for statistical computing and graphics supported by the **R** Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis. As of December 2018, R ranks 16th in the TIOBE index, a measure of popularity of programming languages. The language uses libraries and packages which give the program the required functionality. During our study we widely used packages "tseries", "tsDyn", and "urca" which implement a variety of time series analysis techniques such as unit root and cointegration tests, "exuber" which is used for testing for and dating periods of explosive dynamics and the "MonteCarlo" package which is used for automatically setting up loops to run over parameter grids.

The rest of the paper proceeds as follows. Chapter 2 reviews the existing literature; Chapter 3 contains the most important points of the theory, one should be familiar with in order to better comprehend this study; Chapter 4 discusses the different Monte Carlo scenarios we examine; Chapter 5 provides the results; Chapter 6 illustrates two empirical examples and Chapter 7 concludes.

Chapter 2

Literature Review

The effects of increasing the frequency of observations and the data span on both unit roots and cointegration tests is a subject of discussion among researchers. The Monte Carlo experiments played a significant role in this search. Based on Monte Carlo simulations, Shiller and Perron (1985) argue that the power of unit root tests depends solely on the total sample length. Hooker (1993) examined the properties of the Engle and Granger (1987) cointegration method and showed that contrary to unit root tests, the Engle-Granger test gains significant power from temporal disaggregation. The findings of Shiller and Perron (1985) were supported by Lahiri and Mamingi (1995) who showed that when both the data length and the frequency varies the power of Engle-Granger depends more on the total sample length. The Johansen (1988) test was later examined by Otero and Smith (2000) who showed that the ability of the test depends more on the sample length than the number of observations. More recently, Zhou (2001) showed that when the studies are restricted by relatively short time spans of 30 to 50 years, increasing data frequency may yield considerable power gain and less size distortion, especially when the cointegrating residual is not nearly non-statianory.

Other studies approach the problem using real data. Bagnai (2008) support that a large span of data is required to confirm Thirlwall's hypothesis. while Narayan and Sharma (2015) investigate the importance of high frequency for the impact of forward premium on spot exchange rate.

At the same time, the academic literature on explosive time series has received more and more attention. A number of bubble detecting methods have been proposed. Considering dividend and stock price data Shiller (1980), Blanchard and Watson (1982) and West (1988) assume that inconsistency with the efficient market hypothesis is evidence for the existence of bubbles. Nelson and Plosser (1982) however, support that apparent evidence for bubbles can be reinterpreted in terms of market fundamentals that are unobserved by the researcher.

Diba and Grossman (1988) and Hamilton and Whiteman (1985) recommend another method of testing for rational bubbles by investigating the stationarity properties of asset prices and observable fundamentals. Using standard unit root tests applied to real U.S. Standard and Poor's Composite Stock Price Index data over the period 1871-1986, Diba and Grossman (1988) test levels and differences of stock prices for non-stationarity, finding support in the data for non-stationarity in levels but stationarity in differences. Since differences of an explosive process still manifest explosive characteristics, these findings appear to reject the presence of a market bubble in the data. Although the results were less definitive, further tests by Diba and Grossman (1988) provide confirmation of cointegration between stock prices and dividends over the same period, supporting the conclusion that prices do not diverge from long-run fundamentals and thereby giving additional evidence against bubble behavior.

Evans (1991) shows through simulation methods that non-recursive unit root tests have low power and frequently cannot reject the the null of no explosive behaviour even when present in the data. Nonlinear dynamics, such as those displayed by mildly explosive processes, may lead the standard right-tailed ADF test to findings of spurious stationarity. Intuitively, this is the case because increases followed by downward corrections make the process appear mean-reverting and stationary in finite samples even when it is inherently not.

The situation remains the same when it comes to asset price data. A review by Gürkaynak (2005) offers an insight in the attempts to construct a proper method to detect explosive behaviour. Phillips et al. (2011) proposed the sup ADF (hereafter SADF), a recursive test procedure for testing explosive behaviour. The test is effective when only one explosive episode is contained in the data but has little power when the data contains multiple bubbles with periodically collapsing behaviour. The power loss are due to the he complex nonlinear structure involved in multiple bubble phenomena. Phillips et al. (2013) developed a new recursive procedure, the generalized sup ADF (hereafter GSADF), which is not affected by multiple bubbles as the SADF. Until recently explosive behaviour in economic and financial variables were typically considered a temporary rather than a permanent feature. However the empirical evidence has demonstrated that the that explosive episodes might last even for more than a decade.

For many years the I(2) model was used for cointegration analysis when there were explosive roots in the data. A study by Juselius and Mladenovic (2002) of Yugoslavian hyper-inflation data found that cointegration analysis is a useful econometric tool despite the explosive behaviour of the data and the consequent violation of the usual assumptions to cointegration analysis. Bent (2000) and Bent (2005)showed that the asymptotic results for the Johansen's test are valid for explosive growing variables.

Table 2.1: List of previous studies

Year	Authors	Results
1985	Shiller, Perron	The sample length affects the power of DF test
1993	Hooker	The power of Engle-Granger test increases when the frequency is increased
1995	Lahiri, Mamingi	The data span affects the Engle-Granger test more than the frequency of observations
2000	Otero, Smith	The data span affects the Johansen test more than the frequency of observations
2001	Zhou	When the data span is short, increasing the frequency yields power gains to the Engle-Granger, Horvath-Watson and Johansen tests
1980	Shiller	Empirical study on rational bubbles in the stock market
1982	Blanchard, Watson	Bubbles in asset prices and market bubbles are consistent with rational bubbles
1988	West	Small sample bias in methods for detecting explosive behaviour
1982	Nelson, Plosser	Question the power of previous bubble tests
1985	Hamiliton, Whiteman	Most of the existing tests for bubble detection are not valid
1988	Diba, Grossman	Detecting bubbles based on stationarity in differences
1991	Evans	Low power in non-recursive unit root tests
2000	Bent	Asymptotic results for Johansen test hold for explosive variables
2002	Juselius, Mladenovic	Treating variables as explosives instead of I(2) in cointegration analysis
2005	Bent	Asymptotic results for Johansen test hold for explosive variables
2005	Gurkaynak	Paper report of bubble detecting tests
2011	Phillips, Wu, Yu	The supremum augmented Dickey-Fuller test
2013	Phillips, Shi, Yu	The generalized supremum augmented Dickey-Fuller test

Chapter 3

Theoretical framework

3.1 Cointegration tests

3.1.1 Engle-Granger two-step method

Granger (1983) gave the following definition of cointegration.

Definition 3.1.1 *Let z_t be a vector of variables, the components of the z_t are cointegrated of order (d,b) if*

1. *All component of z_t are $I(d)$.*
2. *There is at least one vector of coefficients α such that $\alpha^T z_t$ is integrated of order $d - b$.*

Engle and Granger (1987) proposed a cointegration test consisting of two steps. To explain the procedure of the test we assume that there are two variables x_t and y_t which both are $I(1)$. In the first step we pretest the variables in order to confirm that they are non-stationary. If both time series are $I(0)$ then the standard regression analysis is valid. A great number of tests have been proposed through the years in order to test for stationarity. These are the augmented Dickey-Fuller test and the Phillips-Perron unit root tests and the KPSS (Kwiatkowski et al. (1992)) stationarity test.

The second step is to estimate the regression

$$y_t = \beta_0 - \beta_1 x_t + \epsilon_t \tag{3.1}$$

using the ordinary least squares (OLS). The predicted residuals $\hat{\epsilon_t}$ are subjected to the ADF test. Since the residuals are themselves estimates, Mackinnon (1990) estimated new corrected critical values. If we confirm the stationarity of the residuals we can form the error correction model (ECM) using them as one variable.

$$\Delta y_t = a\Delta x_t + \rho\hat{\epsilon_t} + \mu + u_t \tag{3.2}$$

Although the Engle-Granger method is easy to apply, there are some drawbacks. First of all, only one cointegrating relationship can be examined at a time. Additionally, we have to treat the variables asymmetrically and specify one dependent variable. The method also suffers from low statistical power of the unit root tests at stage one and from possible small sample bias overall. Finally, the validity of the long-run parameters in the first regression stage where one obtains the residuals cannot be verified because the distribution of the OLS estimator of the cointegrating vector is highly complicated and non-normal. The cointegration procedures proposed later, address most of these problems.

3.1.2 Phillips-Ouliaris cointegration test

Phillips and Ouliaris (1990) showed that residual-based unit root tests applied to the estimated cointegrating residuals, obtained in the first stage of the Engle-Granger method, do not have the usual Dickey–Fuller distributions under the null hypothesis of no-cointegration. Because of the spurious regression phenomenon under the null hypothesis, the distribution of these tests have asymptotic distributions that depend on the number of deterministic trend terms and the number of variables with which cointegration is being tested.

As a solution to this problem they introduced two new residual-based tests, the variance test and the multivariate statistic test. The two unit root tests examine the null hypothesis of no cointegration against the alternative of the presence of cointegration using scalar unit root tests applied to the residuals. The multivariate trace statistics has the advantage over the variance ratio test in that it is invariant to normalization, that is, whichever variable is taken to be the dependent variable, the test will yield the same results, Pfaff (2006).

Both tests are based on the residuals of the first-order vector autoregression

$$z_t = \hat{\Pi} z_{t-1} + \hat{x}_t \tag{3.3}$$

where z_t is partitioned as $z_t = (y_t, x_t^T)$. The variance ratio statistic $\hat{P}_u$ is then defined as

$$\hat{P}_u = \frac{T\hat{\omega}_{11.2}}{T^{-1}\sum_{t=1}^{T}\hat{e}_t^2} \tag{3.4}$$

where $\hat{\epsilon}_t$ are the residuals of the long-run equation (3.1). The conditional covariance $\hat{\omega}_{11.2}$ is derived from the covariance matrix $\hat{\Omega}$ of the residuals ξ_t of the equation (3.3) and is defined as

$$\hat{\omega}_{11.2} = \hat{\omega}_{11} - \hat{\omega}_{21}^T \hat{\Omega}_{22}^{-1} \hat{\omega}_{21} \tag{3.5}$$

where the covariance matrix Ω has been partitioned as

$$\hat{\Omega} = \left[\begin{array}{cc} \hat{\omega}_{11} & \hat{\omega}_{21} \\ \hat{\omega}_{21} & \hat{\Omega}_{22} \end{array} \right| \tag{3.6}$$

and is estimated as

$$\hat{\Omega} = T^{-1}\sum_{t=1}^{T}\hat{\xi}_t^T\hat{\xi}_t + T^{-1}\sum_{s=1}^{l} w_{sl}\sum_{t=1}^{T}(\hat{\xi}_t\hat{\xi}_{t-s}^T + \hat{\xi}_{t-s}\hat{\xi}_t^T) \tag{3.7}$$

with the weighting functions $w_{sl} = 1 - s/(l+1)$. Therefore, the variance ratio statistic measures the size of the residual variance from the cointegrating regression of y_t on x_t against that of the conditional variance of y_t given x_t. In the case of cointegration, the test statistic should stabilize to a constant, whereas if a spurious relationship is present, this would be reflected in a divergent variance of the long-run equation residuals from the conditional variance. Critical values of the test have been tabulated in Phillips and Ouliaris (1990).

The multivariate trace statistic, denoted as $\hat{P}_z$, is defined as

$$\hat{P}_z = Ttr(\hat{\Omega} M_{zz}^{-1}) \tag{3.8}$$

where $M_{zz} = t^{-1} \sum_{t=1}^{T} z_t z_t^T$, and $\hat{\Omega}$ estimated as in Equation (3.7). Critical values for this test statistic are provided in Phillips and Ouliaris (1990), too. The null hypothesis is that no cointegration relationship exists.

3.1.3 Johansen test

The greatest deficiency of the two method we discussed so far, is that we can only detect and estimate a single cointegration relationship at a time. However if we deal with more than two time series, it is possible that more than one cointegrating relationship exists. In fact, between m variables there can be up to $m-1$ cointegrating relationships. The Johansen approach offers a solution to this problem.

Assume a vector autoregressive model (VAR) in it the usual form,

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + ... + \beta_k t_{t-k} + u_t \tag{3.9}$$

where y_t is a (g×1) vector of variables. We transform the VAR to

$$\Delta y_t = \Pi y_{t-k} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + +\Gamma k - 1 \Delta y_{t-(k-1)} + u_t \tag{3.10}$$

where $\Pi = (\sum_{j=1}^{k} \beta_i) - I_g$ and $\Gamma_i = (\sum_{j=1}^{i} \beta_i) - I_g$. Since in the long-run all Δy_{t-i} become 0, Π is the long-run coefficient matrix.

The Johansen procedure focuses on the rank of the Π matrix via its eigenvalues. If the coefficient matrix Π has reduced rank $r < g$, then there exist matrices α and β each with dimensions $g \times r$ and rank r so that, $\Pi = \alpha\beta^T$ and $\beta^T y_t$ is stationary. The number of cointegrating relationships is r, the elements of α are known as the adjustments parameters in the vector error correction model (VECM) and each column of β is a cointegrating vector. For a given r, the maximum likelihood estimator of β defines the combination of y_{t-1} that yields the r largest canonical correlation of Δy_t with y_{t-1} after correcting for lagged differences and deterministic variables when present. Johansen (1988) proposes a two different likelihood ratio tests of significance of these canonical correlations and thereby the reduced rank of the Π matrix, the trace test and maximum eigenvalue test, formulated as

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{g} \ln(1 - \hat{\lambda}_i) \tag{3.11}$$

$$\lambda_{max}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1}) \tag{3.12}$$

where $\hat{\lambda}_i$ is the estimated value for the ith ordered eigenvalue from the Π matrix. λ_{trace} tests the null hypothesis that the number of cointegrating vectors is less than or equal to r against an unspecified alternative, while λ_{max} tests the null that the number of cointegrating vectors r against an alternative of $r+1$. If the matrix Π has zero rank then there is no long-run relationship and if it has full rank, the original variables are stationary.

The distribution of the test statistics is non-standard. The critical values depend on the value of $g - r$, the number of non-stationary components, and whether a constant and/or trend are included in the regressions. Johansen (1988) were the first to tabulate critical values for the Johansen test. As the test gained popularity and the asymptotic and small properties have been revised, new critical values were tabulated by Osterwald-Lenum (1992), Doornik (1998) and MacKinnon et al. (1999). In this study we use the critical values proposed by MacKinnon et al. (1999).

3.2 Tests for explosive behaviour

3.2.1 Supremum Augmented Dickey-Fuller test

In order to describe the recursive implementation of the SADF and GSDAF tests, some notation is required. Specifically, in the analysis of explosive behaviour in time series the full sample is normalised on the interval $[0, 1]$ (i.e., divided by the number of observations T). We denote r_1 and r_2 the corresponding fractions of the sample which define the beginning and end of a subsample such that $0 \leq r_1 \leq r_2 \leq 1$. We denote by $r_w = r_2 - r_1$ the window size of regression estimation, while r_0 is the fixed initial window required by the econometrician such that the subsample in r_2 satisfies that $r_2 \in [r_0, 1]$

The SADF test, proposed by Phillips et al. (2011) is a recursive procedure based on the recursive estimation of the ADF regression on subsamples of the data. The approach uses a forward expanding estimation subsample with the end of the subsample r_2 increasing from $r_0 \in (0, 1)$ (the fixed minimum size for the initial window) to one (the last available observation). The starting point of each estimation is kept fixed at $r_1 = 0$, so the expanding window size of the regression is simply given by $r_w = r_2$. Then, incrementing the window size $r_2 \in [r_0, 1]$ with one additional observation at a time, the recursive estimation of the regression equation

$$\Delta y_t = a_{r_1,r_2} + \beta_{r_1,r_2} y_{t-1} + \sum_{j=1}^{k} \psi^j_{r_1,r_2} \Delta y_{t-j} + \epsilon_t \tag{3.13}$$

over the forward expanding subsample yields a sequence statistics,

$$ADF_0^{r_2} = \frac{\hat{\beta}_{0,r_2}}{s.e.(\hat{\beta}_{0,r_2})}$$

The Phillips et al. (2011) test statistic, is defined as the supremum value of the sequence of $ADF_0^{r_2}$ statistics expressed as follows,

$$SADF(r_0) = \sup_{r_2 \in [r_0,1]} ADF_0^{r_2} \tag{3.14}$$

The minimum window size is set to $0.01 + 1.8/\sqrt{T}$. Whenever $SADF(r_0)$ exceeds the corresponding right-tailed critical value from its limit distribution, the unit root hypothesis is rejected in favour of mildly explosive behaviour.

The rolling-window structure of the $SADF(r_0)$ test leads to improved power in detecting mildly explosive behaviour relative to what can be achieved with a standard ADF_0^1 test alone. Furthermore, Homm and Breitung (2011) show through simulation experiments that the $SADF(r_0)$ test generally outperforms alternative testing methods commonly used to detect a single structural break in the persistence of the process from $I(1)$ to explosive as well. Phillips and Yu (2011) modified the technique and provided a technology for identifying bubble behavior with consistent dating of their origination and collapse.

The alternative tests considered by Homm and Breitung (2011) aim to detect a permanent structural break in the persistence of the process and, as a consequence, perform well only when the series becomes explosive but never bursts in-sample. Intuitively, the $SADF(r_0)$ test's power and its performance deteriorate in the presence of recurring (more than one) and periodically collapsing episodes of exuberance, as established in Phillips et al. (2013)

3.2.2 Generalized supremum ADF

Phillips et al. (2013) proposed another recursive (right-tailed) unit root test, the Generalized SADF (GSADF), covering a larger number of subsamples than the $SADF(r_0)$ test by relaxing the requirement that the starting point of the subsample r_1 be kept fixed. This additional margin of flexibility on the estimation window of the $GSADF(r_0)$ results in substantial power gains, consistent with multiple and periodically collapsing episodes of explosiveness in the data (while the $SADF(r_0)$ test is only consistent with a single such episode in in-sample).

The GSADF approach builds on the forward expanding estimation subsample strategy of the SADF procedure, but instead allows the starting point of the subsample r_1 to change. The initial window size r_0 satisfies that $r_0 < r_2$, while the expanding windows size of the regression (over the normalized sample) is defined as $r_w = r_2 - r_1$. Incrementing the window size $r_2 \in [r_0, 1]$ with one additional observation at a time over each starting point of the sample $r_1 \in [o, r_2 - r_0]$, the recursive estimation of the ADF regression equation in (3.13) yields a sequence of statistics,

$$ADF_{r_1}^{r_2} = \frac{\hat{\beta}_{r_1,r_2}}{s.e.(\hat{\beta}_{r_1,r_2})}$$

The Phillips et al. (2013) test statistic is defined as the supremum value of the sequence $ADF_{r_1}^{r_2}$ statistics expressed as follows,

$$GSADF(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \left\{ \sup_{r_2 \in [r_0, 1]} ADF_{r_1}^{r_2} \right\} \tag{3.15}$$

Whenever $GSADF(r_0)$ exceeds the corresponding right-tailed critical value from its limit distribution, the unit root hypothesis is rejected in favour of mildly explosive behaviour. The rolling window structure of the $GSADF(r_0)$ test leads to improved power in detecting recurring episodes of mildly explosive behaviour relative to what can be achieved with the standard ADF_0^1 and with the $SADF(r_0)$ test.

The critical values for the both the SADF and the GSADF test are obtained via Monte Carlo simulations.

Chapter 4

Design of the Monte Carlo experiments

Our aim is to study the effects of changing the frequency of the observations and the total sample length on the power of the tests described in the previous chapter. For this purpose, using data generation processes (DGP) we construct systems of two, three and four variables. Each time series has a sample length of 1188 observations. This corresponds to 99 years of monthly data. We create quarterly and annual data using two different methods. The first method keeps only the last observation of every quarter (year). The second method averages the three (twelve) non-overlapping observations corresponding to each quarter (year). We then employ the tests to sample of 33, 66 and 99 years of monthly, quarterly and annual observations. The power of the tests are based on 1000 replications. The procedure we just described is used throughout this study. Unless stated otherwise a significance level of 5% is used. When needed, the Schwarz information criterion (SIC) is employed to determine the VAR's lag order. Koehler and S. Murphree (1988) compared the SIC and the Akaike information criterion (AIC) and concluded that the SIC leads to lower order models and is preferable to apply.

4.1 Cointegration

4.1.1 The Engle-Granger and Phillips-Ouliaris methods

We begin the analysis with the two single-equation methods, the Engle-Granger and the Phillips-Ouliaris. For the first case of two time series. Following Otero and Smith (2000) and construct two time series using the DGP,

$$\begin{aligned} y_t + x_t &= u_{1,t}, \quad u_{1,t} = u_{1,t-1} + error \\ y_t + 2x_t &= u_{2,t} \quad u_{2,t} = \rho u_{2,t-1} + error \end{aligned} \tag{4.1}$$

We examine $\rho = 0.98$, 0.95 0.92, 0.85.

We consider two more systems. The first contains three time series and it is based on the procedure below,

$$\begin{aligned} x_t &= u_{2,t} + u_{3,t} - u_{1,t}, \quad u_{1,t} = u_{1,t-1} + error \\ y_t &= 2(u_{1,t} + u_{3,t}) - u_{2,t}, \quad u_{2,t} = 0.95u_{2,t-1} + error \\ z_t &= u_{2,t} + u_{3,t} - u_{1,t}, \quad u_{3,t} = 0.98u_{3,t-1} + error \end{aligned} \tag{4.2}$$

For the case of four time series we build the following system,

$$\begin{aligned} x_t &= u_{3,t} + u_{4,t} - u_{1,t} - u_{2,t}, & u_{1,t} &= u_{1,t-1} + error \\ y_t &= 2(u_{1,t} + u_{2,t}) - u_{3,t} - u_{4,t}, & u_{2,t} &= 0.95u_{2,t-1} + error \\ z_t &= 2(u_{1,t} + u_{3,t} - u_{2,t}) - u_{4,t}, & u_{3,t} &= 0.98u_{3,t-1} + error \\ w_t &= u_{2,t} + u_{4,t} - u_{1,t} - u_{3,t}, & u_{4,t} &= 0.98u_{4,t-1} + error \end{aligned} \tag{4.3}$$

All the errors are independent standard normal and unless stated otherwise we assume that every disturbance term does so. The first value of the autoregressive processes u_t is set to zero.

The Engle-Granger and the Phillips-Ouliaris tests are then employed to test for cointegration. The Engle-Granger is applied to systems (4.1), (4.2) and (4.3), two, three and four times respectively. Every time the test is applied a different variable is considered as endogenous. This is due to the fact that the Engle-Granger method treats the variables asymmetrically. As regards the Phillips-Ouliaris test, we only use the multivariate trace statistics which is invariant to normalization.

4.1.2 Johansen test

The three systems are also tested for cointegration using the Johansen procedure. In the system (4.2), however, the coefficients of the autoregressive procedures $u_{2,t}$ and $u_{3,t}$ are equal to 0.85 and 0.92 respectively. Actually, during our analysis we first examined the Johansen procedure using the values 0.85 and 0.92. When we applied the Engle-Granger and the Phillips-Ouliaris methods, the probability to detect a cointegration relationship were incredibly high regardless of the sample and temporal disaggregation were hardly observed. Changing the values of the parameters eased our analysis.

The examination of the systems (4.2) and (4.3) with the Johansen test revealed that more than one cointegrating vectors are present (the results are fully presented in the next chapter). This makes us wonder if the effects of changing the frequency of observations and the total sample length on the power of the test is different depends on the number of the existent cointegrating vectors. In order for our study to be complete, we create additional multivariate systems where different numbers of cointegration relationships exist. Systems (4.2) and (4.3) contain two and three cointegrating vectors respectively. We consider the supplementary cases of zero and one vectors for the three time series and the cases of zero, one and two cointegrating vectors for the four time series.

For the case of three variables and no cointegration we examine a system which consists of three random walk processes. To create a system of three variables with one cointegrating relationship we follow the DGP,

$$\begin{aligned} x_t &= u_{2,t} - u_{1,t} \\ y_t &= 2u_{1,t} - u_{2,t} \\ z_t &= z_{t-1} + error \end{aligned} \tag{4.4}$$

where, $u_{1,t} = u_{1,t-1} + error$ and $u_{2,t} = 0.95u_{2,t-1} + error$.

To study the behaviour of the Johansen test in the case of four variables and absence of cointegration we employ the the DGP,

$$
\begin{aligned}
x_t &= x_{t-1} + u_{1,t} \\
y_t &= x_t + u_{2,t} \\
z_t &= z_{t-1} + error \\
w_t &= w_{t-1} + error
\end{aligned}
\tag{4.5}
$$

where $u_{1,t}$ follows a random walk and $u_{2,t} = 0.95u_{1,t-1} + error$.

In the next system exactly one cointegrating vector is present.

$$
\begin{aligned}
y_t + x_t &= u_{1,t} \\
y_t + 2x_t &= u_{2,t} \\
z_t &= z_{t-1} + error \\
w_t &= w_{t-1} + error
\end{aligned}
\tag{4.6}
$$

with $u_{1,t}$ following a random walk and $u_{2,t} = 0.92u_{2,t-1} + error$.

Finally, for the case of two cointegrating vectors we create four time series following the DGP,

$$
\begin{aligned}
x_t &= u_{2,t} + u_{3,t} - u_{1,t}, & u_{1,t} &= u_{1,t-1} + error \\
y_t &= 2(u_{1,t} + u_{3,t}) - u_{2,t}, & u_{2,t} &= 0.95u_{2,t-1} + error \\
z_t &= u_{2,t} + u_{3,t} - u_{1,t}, & u_{3,t} &= 0.98u_{3,t-1} + error \\
w_t &= w_{t-1} + error
\end{aligned}
\tag{4.7}
$$

4.2 Explosive behaviour

The ability of a test two detect explosive behaviour depends on a great degree on the kind of the bubble. For example, the SADF can easily detect a single bubble but it fails to recognize multiple collapsing explosive episodes.

We apply the two right-tailed unit root tests on three different kind models. The first model proposed by Phillips et al. (2011) follows the DGP,

$$
x_t = \begin{cases} 0, & \text{if} \quad t = 0 \\ x_{t-1} + error, & \text{if} \quad t < \tau_e \quad or \quad t > \tau_f \\ \delta_T x_{t-1} + error, & \text{if} \quad \tau_e \leq t \leq \tau_f \end{cases}
\tag{4.8}
$$

where T is the sample size, $\delta_T = 1 + cT^{-a}$ with $c > 0$ and $a \in (0, 1)$. The series follows a pure random walk process except for a bubble period from τ_e to τ_f. During that period it follows a mildly explosive process with expansion rate given by the autoregressive coefficient δ_T. We set $T = 1188$, $c = 1$, $a = 0.6$, $\tau_e = 120$ and $\tau_f = 298$. Given that, the explosive process takes place during the years 10-25.

The second model is based on Blanchard (1979). It consists of two regimes, which occur with probability π and $1 - \pi$. In the first regime, the bubble grows exponentially,

$$
x_t = \frac{1+r}{\pi} + x_{t-1} + error
\tag{4.9}
$$

whereas in the second regime, the bubble collapses to a white noise. We set $\pi = 0.7$ and $r = 0.05$.

The last series follows Evans (1991).

$$x_t = \begin{cases} (1+r)x_{t-1}u_t, & \text{if} \quad x_{t-1} \leq a \\ [\delta + \pi^{-1}(1+r)\theta_t(x_{t-1}) - (1+r)^{-1}\delta]u_t, & \text{if} \quad x_{t-1} > a \end{cases} \tag{4.10}$$

δ and a are positive parameters with $0 < \delta < (1+r)a$, u_t is an exogenous iid positive random variable with $E_{t-1}u_t = 1$, and θ_t is an exogenous independently and identically distributed Bernoulli process (independent of u) which takes the value 1 with probability π and 0 with probability $1-\pi$, where $0 < \pi \leq 1$. When $x_{t-1} \leq a$ the bubble grows at an average rate of $1+r$. When $x_{t-1} > a$ the bubble expands at an increased rate of $(1+r)\pi^{-1}$. We set $a = 1, \delta = 0.5, \pi = 0.7$ and $r = 0.05$. The last two models contain multiple explosive episodes.

Although not our main concern, during the analysis we examine the three models using common unit root tests. We apply the Dickey and Fuller (1981), Kapetanios et al. (2003) (KSS) and Zivot and Andrews (1992) tests on all three time series.

The last part of the simulations deals with cointegration tests on explosive time series. We consider two cases. First, we construct the following time series,

$$\begin{aligned} x_t &= u_{2,t} - u_{1,t} \\ y_t &= \begin{cases} 2u_{1,t} - u_{2,t}, & \text{everywhere} \\ \delta_T x_{t-1} + error, & \text{if} \quad 120 \leq t \leq 298 \end{cases} \end{aligned} \tag{4.11}$$

where,

$$\begin{aligned} u_{1,t} &= u_{1,t-1} + error \\ u_{2,t} &= \rho u_{2,t-1} + error \end{aligned}$$

$\rho = 0.98$, 0.95 0.92, 0.85, $u_{1,0} = u_{2,0} = 0$ and the errors are independent standard normal. Obviously, x_t is integrated of order 1 and y_t contains a single bubble which occur during the years 10-24. One can easily confirm this using common unit root tests.

Second, we consider series which exhibit multiple explosive incidents. We follow the DGP,

$$\begin{aligned} x_t &= u_{2,t} - u_{1,t} \\ y_t &= 2u_{1,t} - u_{2,t} \end{aligned} \tag{4.12}$$

where $u_{1,t}$ is a random walk. Considering $u_{2,t}$ there are two subcases, the first time it follows the process described by Blanchard (1979) and the second time it follows the process described by Evans (1991).

4.3 A note on the lag order

Let us return to system (4.5) where,

$$x_t = x_{t-1} + u_{1,t}, \quad u_{1,t} \sim N(0,1)$$
$$y_t = x_t + u_{2,t}$$

and $u_{2,t}$ is an AR(1) process,

$$u_{2,t} = 0.95u_{2,t-1} + e_t, \quad e_t \sim N(0,1)$$

We denote the quarterly and annual data obtained with the method of skip sampling,

$$x_t^{end} = x_{t \cdot s}$$
$$y_t^{end} = y_{t \cdot s}$$

where s = 3 for quarterly and s = 12 for annual, hence

$$x_t^{end} = x_{t-1}^{end} + u_{1,t}^{end},$$
$$u_{1,t}^{end} = u_{1,t \cdot s-(s-1)} + u_{1,t \cdot s-(s-2)} + ... + u_{1,t \cdot s},$$
$$E[u_{1,t}^{end}, u_{1,t-j}^{end}] = 0 for j \neq 0 \quad \text{and}$$
$$y_t^{end} = x_t^{end} + u_{2,t}^{end},$$
$$u_{2,t}^{end} = 0.95^s u_{2,t-1}^{end} + e_t^{end},$$
$$e_t^{end} = e_t + 0.95e_{t \cdot s-1} + ... + r^{s-1}e_{t \cdot -(s+1)}$$

Because $E[u_{2,t}^{end}, u_{2,t-j}^{end}] = 0$ for $j \neq 0$, $u_{2,t}$ remains an AR(1) process.

We denote the averaged quarterly and annual data as

$$x_t^{av} = \frac{1}{s} \sum_{i=st-(s-1)}^{st} x_i$$
$$y_t^{av} = \frac{1}{s} \sum_{i=st-(s-1)}^{st} y_i$$

x_t^{av} is also $x_{t-1}^{av} = x_{t-1}^{av} + u_{1,t}^{av}$. The error term $u_{1,t}^{av}$ is written,

$$u_{1,t} = \frac{1}{s} \sum_{i=s(t-1)-(s-1)}^{st} x_i$$

By the definition of y_t^{av} we get $y_t^{av} = x_t^{av} + u_{2,t}^{av}$. The error term $u_{1,t}^{av}$ can be expressed by the two equations,

$$u_{2,t}^{av} = \frac{1}{s} \sum_{i=st-(s-1)}^{st} u_{2,i}$$
$$u_{2,t}^{av} = 0.95^s \ u_{2,t-1}^{av} + e_t^{av}$$

Combining the two equations gives us,

$$e_t^{av} = \frac{1}{s}\Bigg\{ \sum_{i=st-(s-1)}^{st} \Bigg[e_i(\sum_{j=0}^{s-i} 0.95^j) \Bigg] + \sum_{i=s(t-(s-1)}^{s(t-1)} \Bigg[e_i(\sum_{j=s-i+1}^{s-1} 0.95^j) \Bigg] \Bigg\}$$

It is shown that $E[u_{1,t}^{av}, u_{1,t-1}^{av}] \neq 0$ and $E[e_t^{av}, u_{t-1}^{av}] \neq 0$ yet $E[u_{1,t}^{av}, u_{1,t-1}^{av}] = 0$ and $E[e_t^{av}, u_{t-1}^{av}] = 0$. This means that $u_{1,t}^{av}$ and e_t^{av} are moving average processes (MA(1)) and thus could be expressed in a typical MA(1) form,

$$u_{1,t} = w_t - \theta w_{t-1} = (1 - \theta L) w_t$$
$$e_t = v_t - \phi v_{t-1} = (1 - \phi L) v_t$$

where $E[w_t, w_{t-j}] = 0$ and $E[v_t, v_t - j] = 0$ for $j \neq 0$, θ and ϕ are the moving average parameters and L is the lag operator. Because the first order autocorrelation of an MA(1) process is equal to $-\theta/(1+\theta^2)$, the approximate values of θ and ϕ can be computed by solving the following

$$\frac{E[u_{1,t}^{av}, e_{1,t-1}^{av}]}{E[(u_{1,t}^{av})^2]} = \frac{\sum_{i=2}^{s} \Bigg[\Bigg(\sum_{j=0}^{s-i} \rho^j \Bigg) \Bigg(\sum_{j=s-i+1}^{s-1} \rho^j \Bigg) \Bigg]}{\sum_{i=1}^{s} \Bigg[\Bigg(\sum_{j=0}^{s-i} \rho^j \Bigg)^2 \Bigg] + \Bigg[\Bigg(\sum_{j=s-i+1}^{s-1} \rho^j \Bigg)^2 \Bigg]} = \frac{-\theta}{1+\theta^2}$$

$$\frac{E[e_t^{av}, e_{t-1}^{av}]}{E[(e_t^{av})^2]} = \frac{\sum_{i=2}^{s}[(s+1-i)(i-1)]}{\sum_{i=2}^{s}(s+1-i)^2 + \sum_{i=2}^{s}(i-1)^2} = \frac{-\phi}{1+\phi^2}$$

Note that θ is a function of ρ and ϕ can be considered as a special case of θ for $\rho = 1$. As the models of cointegration tests are mostly presented in an autoregressive or VAR form, the existence of an MA(1) tern in x_t^{av} and $u_{2,t}^{av}$ may require may require an infinite lag structure. THat is, when

$$x_t^{av} = x_{t-1}^{av} + (1 - \theta L) w_t$$
$$u_{2,t}^{av} = 0.95^s + u_{2,t-1}^{av} + (1 - \phi L) v_t$$

Denoting $\Delta x_t^{av} = x_t^{av} - x_{t-1}^{av}$ and $z_t = u_{2,t}^{av} - 0.95^s + u_{w,t-1}^{av}$, we have

$$\Delta x_t^{av}/(1 - \theta L) = \Delta x_t^{av} - \theta \Delta x_{t-1}^{av} + \theta^2 \Delta x_{t-2}^{av} - \theta^3 \Delta x_{t-3}^{av} + ... = w_t$$
$$w_t/(1 - \phi L) = w_t - \phi w_{t-1} + \phi^2 w_{t-2} - \phi^3 w_{t-3} + ... = v_t$$

Δx_t^{av} and z_t are AR processes with an infinite lag structure.

In practice, if the MA coefficient *theta* is relatively small, the study would not be hurt by the problem of underparameterization if we use the models with finite but sufficient lag lengths. As we have already said, we use the SIC for the selection of the lag order. Supplementary simulations, not presented here, showed that although the AIC often results in higher order VARs, the results from the analysis are qualitatively the same.

Chapter 5

Results

5.1 Single-equation cointegration tests

We employ the Engle-Granger and the Phillips-Ouliaris tests on the time series produced be equations (4.1), (4.2), (4.3). Tables 5.1, 5.2 and 5.3 present the results. Because the two tests differ only in the second step they exhibit similar behaviour.

Tables 5.1 and 5.2 report the results from the the two tests applied on (4.1). We first observe that the value of ρ affects the power of the tests. As ρ approaches 1 the performance of the tests worsens. Both the frequency of observations and the total sample length play a role on the performance of the tests. Increasing the frequency of observations yields substantial power gains. For example, for $\rho = 0.85$, averaged data and $p = 0.05$, using quarterly instead of annual data triples the probability the Engle-Granger method to detect a cointegrating vector. In general setting ρ equal to 0.85 produces results which do not comply with the results we obtain for the rest of the values of ρ. As ρ approaches one, the ability of the two tests depends more on the total sample length. For instance, for $\rho = 0.98$, using annual data for 66 years yields more power to the Engle-Granger test than using monthly data for 33 years. We should keep in mind that a sample of 33 years of monthly data contains 396 observations while a sample of 66 years of annual data only 66.

Applying the Phillips-Ouliaris method produces similar results. Even when ρ is equal to 0.85 increasing the sample length while using annual data yields great power gains. As ρ grows the importance of the total sample length becomes more visible. We should also note that for most cases we obtain better results using the Phillips-Ouliaris test than the Engle-Granger.

Table 5.3 consists of four subtables. The upper left subtable contains the results from the Engle-Granger applied on (4.2). Using quarterly over annual data yields less power gains than increasing the total data length. The only case where this does not hold is when using quarterly data for 66 years. Using monthly over quarterly increases the probability of detecting a cointegrating relationship. However this increase can be considered insignificant compared to the increase in the power from using a longer data set.

The upper right part of table 5.3 reports the from the Engle-Granger applied on (4.3). Using a set of 99 years give us a probability of detecting cointegration over 85% regardless of the frequency. When we use a sample of 66 years, lowering the frequency yields to substantial power loss. However there is greater loss when we use a sample of monthly observations of 33 years than a sample of annual observations

Table 5.1: Empirical power of Engle-Granger method

$\rho = 0.85$	33 years		66 years		99 years	
Monthly	0.894	0.823	0.886	0.793	0.888	0.810
Quarterly skip	0.745	0.808	0.925	0.860	0.934	0.887
Annual skip	0.211	0.317	0.536	0.671	0.879	0.876
Quarterly avg	0.727	0.815	0.931	0.859	0.931	0.882
Annual avg	0.238	0.353	0.549	0.674	0.873	0.879
$\rho = 0.92$	33 years		66 years		99 years	
Monthly	0.639	0.699	0.862	0.767	0.852	0.758
Quarterly skip	0.427	0.556	0.879	0.819	0.918	0.826
Annual skip	0.128	0.240	0.393	0.540	0.815	0.847
Quarterly avg	0.410	0.556	0.876	0.835	0.920	0.835
Annual avg	0.142	0.234	0.412	0.574	0.807	0.835
$\rho = 0.95$	33 years		66 years		99 years	
Monthly	0.330	0.454	0.793	0.753	0.862	0.755
Quarterly skip	0.229	0.352	0.716	0.750	0.887	0.822
Annual skip	0.104	0.188	0.283	0.419	0.660	0.758
Quarterly avg	0.242	0.352	0.703	0.749	0.888	0.827
Annual avg	0.126	0.193	0.297	0.417	0.655	0.755
$\rho = 0.98$	33 years		66 years		99 years	
Monthly	0.100	0.143	0.242	0.368	0.529	0.617
Quarterly skip	0.083	0.133	0.197	0.317	0.423	0.557
Annual skip	0.063	0.127	0.148	0.216	0.272	0.429
Quarterly avg	0.082	0.157	0.189	0.314	0.422	0.557
Annual avg	0.074	0.129	0.139	0.215	0.267	0.406

Every first column refers to 5% significance level and every second column to 10% significance level.

for 66 years, over a sample of monthly data for 66 years.

In the lower subtable of 5.3 the results from the Phillips-Ouliaris test are reported. When employed on three series, the probability to identify a cointegrating relationship using a sample of 99 years is one regardless of the frequency (except for annual data obtained with the averaging with non-overlapping method). The explanation for this value is that systems in the systems (4.2) and (4.3) exist more than one cointegrating relationships. This will become clearer when we will repeat the analysis using the Johansen test. Because we use the multivariate trace statistic test which is invariant to normalization the probability of detecting a cointegrating vector is much higher than using the Engle-Granger test. However this test is greatly affected when the smallest sample is used. Especially when the averaging with non-overlapping observations method is used, the test always fails to detect a cointegrating vector. Looking at the probabilities in the remaining cases we can conclude that the ability of the test to identify a cointegrating vector depends more on the total sample length. Finally, is worth mentioning that using the skip sampling method produces better results than the averaging method.

Table 5.2: Empirical power of Phillips-Ouliaris test

$\rho = 0.85$	33 years		66 years		99 years	
Monthly	0.894	0.807	0.893	0.808	0.906	0.830
Quarterly skip	0.939	0.873	0.940	0.873	0.945	0.898
Annual skip	0.929	0.901	0.949	0.898	0.942	0.895
Quarterly avg	0.937	0.872	0.945	0.880	0.943	0.897
Annual avg	0.798	0.854	0.959	0.910	0.949	0.895
$\rho = 0.92$	33 years		66 years		99 years	
Monthly	0.843	0.788	0.874	0.778	0.862	0.762
Quarterly skip	0.860	0.823	0.916	0.847	0.908	0.844
Annual skip	0.715	0.799	0.941	0.880	0.945	0.888
Quarterly avg	0.718	0.780	0.924	0.851	0.913	0.844
Annual avg	0.343	0.521	0.929	0.878	0.949	0.890
$\rho = 0.95$	33 years		66 years		99 years	
Monthly	0.524	0.624	0.861	0.764	0.835	0.732
Quarterly skip	0.513	0.636	0.897	0.810	0.896	0.808
Annual skip	0.409	0.549	0.910	0.885	0.921	0.858
Quarterly avg	0.324	0.473	0.862	0.812	0.903	0.811
Annual avg	0.126	0.260	0.743	0.807	0.917	0.858
$\rho = 0.98$	33 years		66 years		99 years	
Monthly	0.124	0.199	0.362	0.478	0.654	0.708
Quarterly skip	0.119	0.210	0.355	0.464	0.650	0.713
Annual skip	0.123	0.216	0.340	0.459	0.613	0.707
Quarterly avg	0.053	0.123	0.242	0.358	0.527	0.637
Annual avg	0.017	0.048	0.144	0.259	0.364	0.515

Every first column refers to 5% significance level and every second column to 10% significance level.

Table 5.3: Testing for cointegration multiple variables with single-equation methods

	33 years	66 years	99 years	33 years	66 years	99 years
	three variables			four variables		
The *Engle − Granger* two-step procedure						
Monthly	0.196	0.637	0.792	0.521	0.708	0.996
Quarterly skip	0.179	0.534	0.789	0.481	0.599	0.956
Annual skip	0.023	0.231	0.523	0.065	0.522	0.856
Quarterly avg	0.176	0.548	0.776	0.492	0.612	0.984
Annual avg	0.030	0.239	0.518	0.052	0.503	0.835
The *Phillips − Ouliaris* test						
Monthly	0.710	1.000	1.000	0.312	0.876	0.997
Quarterly skip	0.545	0.999	1.000	0.136	0.790	0.994
Annual skip	0.004	0.844	1.000	0.001	0.132	0.641
Quarterly avg	0.286	0.989	1.000	0.010	0.471	0.928
Annual avg	0.000	0.384	0.983	0.000	0.016	0.204

The upper and lower subtables report the probability of detecting a cointegrating vector using the Engle-Granger and Phillips-Ouliaris methods respectively, in systems (4.2) (left) and (4.3) (right).

5.2 Multivariate cointegrating test

In this section we present the results from the Johansen test. Table 5.4 reports the probability of identifying a cointegrating relationship between two series using the λ_{max} test. There is little power gain from using monthly data over quarterly data. However, the use of annual data causes substantial power loss. As the value of approaches one the power gains from using a longer sample length become greater. In most cases, the observed power gains to the higher frequency data, compared to annual data can be produced by increasing the sample length by a few years. For example, for $\rho = 0.95$, using the λ_{max} test statistic we need less than 40 years to yield similar power results as 33 years of monthly data. Since we obtain similar results from using the $Trace$ statistic, we will not report them here. We must point out, however, that the $Trace$ statistic is less powerful than the λ_{max}.

Table 5.4: Empirical power of Johansen test

	λ_{max} test			$Trace$ test		
$\rho = 0.85$	33 years	66 years	99 years	33 years	66 years	99 years
Monthly	0.948	0.939	0.941	0.958	0.934	0.946
Quarterly skip	0.937	0.947	0.944	0.954	0.935	0.943
Annual skip	0.537	0.643	0.941	0.524	0.653	0.940
Quarterly avg	0.930	0.944	0.942	0.931	0.931	0.940
Annual avg	0.498	0.920	0.949	0.452	0.896	0.928
$\rho = 0.92$	33 years	66 years	99 years	33 years	66 years	99 years
Monthly	0.761	0.935	0.936	0.781	0.948	0.943
Quarterly skip	0.678	0.933	0.936	0.697	0.952	0.940
Annual skip	0.353	0.526	0.944	0.362	0.488	0.949
Quarterly avg	0.692	0.930	0.944	0.707	0.948	0.950
Annual avg	0.381	0.830	0.897	0.355	0.826	0.899
$\rho = 0.96$	33 years	66 years	99 years	33 years	66 years	99 years
Monthly	0.386	0.909	0.950	0.373	0.898	0.956
Quarterly skip	0.340	0.852	0.951	0.311	0.859	0.954
Annual skip	0.219	0.383	0.916	0.219	0.367	0.919
Quarterly avg	0.289	0.819	0.952	0.358	0.820	0.950
Annual avg	0.244	0.642	0.756	0.254	0.665	0.734
$\rho = 0.98$	33 years	66 years	99 years	33 years	66 years	99 years
Monthly	0.096	0.258	0.550	0.088	0.239	0.544
Quarterly skip	0.098	0.228	0.531	0.099	0.220	0.513
Annual skip	0.113	0.171	0.413	0.111	0.175	0.419
Quarterly avg	0.130	0.214	0.445	0.116	0.208	0.464
Annual avg	0.145	0.248	0.294	0.136	0.245	0.303

The left subtable presents the results from the λ_{max} test and the right subtable the results from the $Trace$ test. The same structure is also used for Tables 5.5 and 5.6.

When three or four variables are tested the power of the tests depends almost solely on the data length. For a fixed number of years changing the frequency has little effect on the power of the test. In several cases, i.e, for the case of four variables and two cointegrating vectors, using a sample of 99 years over a sample of 66 years also has little effect on the test, especially for the monthly and quarterly

Table 5.5: Empirical power of Johansen test applied on three variables

	λ_{max} test			$Trace$ test		
	33 years	66 years	99 years	33 years	66 years	99 years
No cointegration:						
Monthly	0.928	0.936	0.945	0.936	0.941	0.954
Quarterly skip	0.900	0.929	0.944	0.909	0.932	0.943
Annual skip	0.877	0.904	0.937	0.857	0.907	0.933
Quarterly avg	0.920	0.830	0.941	0.925	0.930	0.951
Annual avg	0.806	0.870	0.917	0.791	0.887	0.913
One cointegrating vector:						
Monthly	0.237	0.758	0.949	0.236	0.681	0.926
Quarterly skip	0.221	0.699	0.943	0.219	0.606	0.910
Annual skip	0.224	0.445	0.779	0.255	0.407	0.720
Quarterly avg	0.197	0.551	0.883	0.231	0.494	0.815
Annual avg	0.286	0.457	0.764	0.322	0.448	0.711
Two cointegrating vectors:						
Monthly	0.720	0.936	0.933	0.776	0.936	0.933
Quarterly skip	0.570	0.938	0.938	0.617	0.941	0.938
Annual skip	0.177	0.770	0.930	0.113	0.809	0.930
Quarterly avg	0.611	0.932	0.928	0.627	0.932	0.928
Annual avg	0.160	0.714	0.918	0.109	0.745	0.925

The first subtable reports the prop ability of not detecting a cointegrating vector in system where no cointegration exists. The second and third subtable report the probability of identifying one and two vectors in systems where exactly one and two vectors exist.

observations. However, using the shortest sample length of 33 years yields important power loss. This more clearly observed when at least one cointegration relationship is present.

Although not the main concern of our study, we try to compare the power of the power of the two statistics. When applied on three time series the $Trace$ statistic behaves better when no cointegrating vectors exist and monthly and quarterly data are used. When one cointegrating relationship is present the λ_{max} test produces stronger results using monthly data, quarterly data obtained with the method of the skip sampling and for a data length of 66 years. Of course we can not pronounce over a test based on these results. In the presence of two cointegrating vectors, the λ_{max} produces better results only for the smallest sample of annual data for 33 years. For the largest samples, monthly data for 99 and 66 years and quarterly data for 99 years (1188, 792 and 396 observations respectively) the two tests identify the cointegrating vectors with the exact same probability.

Considering the case of four variables, when no cointegration relationship exists the λ_{max} test is preferable while when one cointegrating vectors exist there is a higher probability that the $Trace$ test can detect it. Finally, when two and three cointegrating vectors exist we could possibly conclude that λ_{max} outperforms the $Trace$ test on longer data sets but it is outperformed on the shortest sample length of 33 years. Again we should notice that for the biggest data sets no test can clearly produce stronger results.

Obviously, we can not safely pronounce over the power of one test over the other.

However it is safe to conclude that in the presence of multiple cointegrating vectors the λ_{max} is preferably used for longer data sets while the $Trace$ test for smaller samples.

Table 5.6: Empirical power of Johansen test applied on four variables

	λ_{max} test			$Trace$ test		
	33 years	66 years	99 years	33 years	66 years	99 years
No cointegration:						
Monthly	0.874	0.898	0.915	0.819	0.880	0.900
Quarterly skip	0.790	0.876	0.903	0.727	0.844	0.879
Annual skip	0.638	0.829	0.861	0.494	0.755	0.820
Quarterly avg	0.753	0.859	0.894	0.655	0.839	0.879
Annual avg	0.533	0.748	0.775	0.355	0.655	0.736
One cointegrating vector:						
Monthly	0.159	0.514	0.664	0.235	0.585	0.681
Quarterly skip	0.156	0.473	0.676	0.225	0.555	0.707
Annual skip	0.243	0.328	0.546	0.361	0.431	0.651
Quarterly avg	0.195	0.523	0.617	0.264	0.602	0.635
Annual avg	0.315	0.357	0.538	0.404	0.479	0.637
Two cointegrating vectors:						
Monthly	0.434	0.936	0.948	0.402	0.901	0.949
Quarterly skip	0.317	0.933	0.946	0.342	0.903	0.933
Annual skip	0.250	0.552	0.922	0.315	0.512	0.871
Quarterly avg	0.348	0.887	0.947	0.370	0.814	0.946
Annual avg	0.330	0.471	0.894	0.382	0.450	0.818
Three cointegrating vectors:						
Monthly	0.211	0.924	0.946	0.274	0.897	0.946
Quarterly skip	0.120	0.876	0.947	0.181	0.834	0.946
Annual skip	0.010	0.400	0.897	0.041	0.447	0.876
Quarterly avg	0.134	0.879	0.943	0.212	0.851	0.941
Annual avg	0.008	0.296	0.855	0.041	0.401	0.838

See notes on Table 5.5.

5.3 Testing for explosive behaviour

From now on our analysis focuses on explosive behaviour. Table 5.7 reports the probability of detecting explosive behaviour using the SADF and GSADF test. As we have already mentioned, we consider three models proposed by Phillips et al. (2013), Blanchard (1979) and Evans (1991) respectively. Table 5.6 sums up our findings. Both tests detect the bubble in the Phillips' model with a rate of success higher than 90%. Therefore, we face difficulty in deciding which characteristic affects the tests more. However, contrary to previous findings, increasing the sample length yields power loss. Since the model proposed by Phillips et al. (2013) contains a single bubble the two tests demonstrate equal power.

The superiority of the GSADF test becomes obvious when we employ the two tests to detect the presence of explosive behaviour in models proposed by Blan-

Table 5.7: Empirical power of SADF and GSADF

	SADF			*GSADF*		
	33 years	66 years	99 years	33 years	66 years	99 years
DGP based on *Phillips*						
Monthly	0.908	0.908	0.905	0.904	0.902	0.898
Quarterly skip	0.910	0.906	0.901	0.905	0.907	0.899
Annual skip	0.901	0.900	0.896	0.907	0.900	0.894
Quarterly avg	0.925	0.925	0.924	0.924	0.926	0.927
Annual avg	0.915	0.916	0.917	0.922	0.920	0.917
DGP based on *Blanchard*						
Monthly	0.507	0.707	0.772	0.993	1.000	1.000
Quarterly skip	0.416	0.528	0.589	0.818	0.928	0.972
Annual skip	0.058	0.059	0.078	0.098	0.107	0.139
Quarterly avg	0.419	0.536	0.614	0.834	0.947	0.977
Annual avg	0.128	0.179	0.210	0.222	0.325	0.394
DGP based on *Evans*						
Monthly	0.828	0.843	0.870	0.992	1.000	1.000
Quarterly skip	0.532	0.605	0.644	0.782	0.922	0.957
Annual skip	0.023	0.061	0.060	0.026	0.061	0.080
Quarterly avg	0.517	0.940	0.684	0.815	0.940	0.971
Annual avg	0.145	0.188	0.220	0.190	0.296	0.376

The table reports the probability of rejecting the null of the SADF (left) and (GSADF) (right) test in favour of the alternative of a mildly explosive process.

chard (1979) and Evans (1991). When describing the two models in section 4.3 we highlighted their main feature which is the multiple periodically collapsing explosive incidents. We can safely deduce from the results that the ability of the two right-tailed unit root tests depends more on the frequency of observations. Regardless of the sample length using annual instead of quarterly data yields significant power losses. To get a better understanding of the effect of changing the frequency, if we apply the SADF test on series generated following Evans' model, we get a higher probability to detect a bubble using monthly data for 33 years than using quarterly data for 99 years.

The two models exhibit similarities, the different expansion rates of the bubbles cause the tests to react differently. The GSADF, for example, produces better results when applied on Blanchard's model. The SADF on the other hand behaves better when applied on Evans' model.

During the analysis we examined the the three models using standard left-tailed unit root tests. We briefly present and discuss some of our findings. Starting with the most common one, the ADF test detects a unit root in all models with a probability higher than 90% for all samples.

The next unit root test to be examined is the KSS. The KSS test examines the hypothesis of a linear time series with a unit root under the alternative of a nonlinear stationary process. When applied on an explosive process it pronounces over the nonlinear stationarity.

For a series generated by equation (4.9) and using a sample of monthly observations for 33 years the test recognizes a stationary time series with probability 83%.

Table 5.8: Probability of KSS to reject H_0

	33 years	66 years	99 years
Monthly	0.830	0.810	0.737
Quarterly skip	0.820	0.802	0.709
Annual skip	0.842	0.774	0.642
Quarterly avg	0.821	0.800	0.710
Annual avg	0.852	0.753	0.627
Monthly	0.992	0.996	0.999
Quarterly skip	0.985	0.996	1.000
Annual skip	0.432	0.647	0.694
Quarterly avg	0.958	0.984	0.992
Annual avg	0.537	0.761	0.819
Monthly	0.996	0.998	0.999
Quarterly skip	0.995	0.997	0.999
Annual skip	0.378	0.644	0.694
Quarterly avg	0.971	0.980	0.996
Annual avg	0.617	0.817	0.842

The probability falls both when we increase the data size and when we decrease the frequency of observations. The importance of the data span over the frequency is more clearly visible when we use a sample of 99 over a one of 66 years. Regardless of the frequency the test suffers greater loss when we use the longest sample.

The series generated by equations (4.10) and (4.11) produce similar results. There is little power gain from using monthly over quarterly data and even less gain from using a longer sample. Using annual data yields subsequent power losses, especially when we use a period of 33 years and the method of skip sampling. When annual data is used there are also notable gains in the power of the test when we increase the data length.

The last unit root test we examine is the Zivot-Andrews. The test has a null hypothesis of a unit root with structural break in the intercept. There are three alternatives depending on the model we choose. These alternatives are a trend stationary process that allows for a one time break in the level, the trend or both. We examined all the cases for three models but we only present the results regarding the series generated by equation (4.9) (Table 5.9). The reason is that the test failed to reject the unit root hypothesis for a time series proposed by Evans (1991) and produced the exact same probability (approximately 0.7) for the model proposed by Blanchard (1979).

When the alternative is stationarity with a break in the trend the test almost always fails to reject the null hypothesis for a sample of 33 years. Using a 66 years sample it is more likely that we reject the initial hypothesis. For that sample a using a higher frequency yields power gains. However the gains are greater when a the longest sample is used. For example using 99 years of quarterly data increases the probability of rejecting H_0 by 10% while using monthly data instead of quarterly only increases the probability by 1%.

The second case examines the presence of the a unit root under the alternative of a stationary series with a break in the level. The test more often rejects the null hypothesis. Again the ability of the test is affected more by the sample length

Table 5.9: Zivot-Andrews testing for statinarity

	33 years	66 years	99 years
with break in the trend:			
Monthly	0.005	0.666	0.762
Quarterly skip	0.009	0.651	0.760
Annual skip	0.014	0.564	0.762
Quarterly avg	0.003	0.703	0.775
Annual avg	0.008	0.624	0.767
with break in the level:			
Monthly	0.674	0.709	0.771
Quarterly skip	0.578	0.669	0.763
Annual skipp	0.032	0.253	0.733
Quarterly avg	0.647	0.717	0.782
Annual avg	0.045	0.350	0.752
with break in the trend and in the level:			
Monthly	0.003	0.747	0.780
Quarterly skip	0.005	0.736	0.775
Annual skip	0.007	0.699	0.769
Quarterly avg	0.002	0.757	0.785
Annual avg	0.003	0.721	0.775

than the frequency. That said we should not underestimate the importance of the frequency. Using quarterly data yields substantial power gain which in the case of 33 years overcome the gain from an increased sample length.

The alternative hypothesis of the third case allows for a break both in the trend and in the level. The results are similar to the case. For the shortest sample the test identifies a unit root with a 99% probability. For the two remaining sample periods the test treats the series a stationary. The performance of the test is depends more on the total sample than the frequency of observations.

5.4 Explosive behaviour in cointegration tests

In the last part of the simulations examine the behaviour of the Johansen and the Engle-Granger cointegration methods when they deal with variables that exhibit explosive behaviour.

We begin with the system generated by (4.12). Table 5.10 contains the results from the two cointegration tests. Considering the Johansen test, only the results from the λ_{max} statistic are reported, since the $Trace$ test produces similar results. Although both tests fail to identify a cointegration relationship most of the times, the effect of changing the sample length is still apparent. Using a data set of 99 years over a data set of 66 years doubles the probability of detecting a cointegrating vector. As before, the power of both tests falls as the value of ρ approaches one. Comparing the performance of the two tests, we see that the Engle-Granger method accepts a cointegration relationship for a sample of 99 years with a higher probability, but for the two shortest samples, it fails to outperform the Johansen test.

Table 5.11 consists of two smaller subtables. The upper subtable contains the

Table 5.10: Cointegration analysis of system 4.12

	Johansen			*Engle − Granger*		
	33 years	66 years	99 years	33 years	66 years	99 years
$\rho = 0.85$						
Monthly	0.079	0.181	0.326	0.019	0.100	0.471
Quarterly skip	0.058	0.141	0.296	0.011	0.082	0.463
Annual skip	0.051	0.103	0.233	0.010	0.048	0.449
Quarterly avg	0.080	0.202	0.572	0.015	0.070	0.458
Annual avg	0.051	0.105	0.234	0.010	0.053	0.492
$\rho = 0.92$						
Monthly	0.064	0.155	0.321	0.016	0.107	0.450
Quarterly skip	0.053	0.142	0.291	0.010	0.098	0.452
Annual skip	0.050	0.116	0.257	0.012	0.065	0.437
Quarterly avg	0.071	0.202	0.578	0.014	0.092	0.465
Annual avg	0.076	0.108	0.248	0.014	0.085	0.496
$\rho = 0.95$						
Monthly	0.046	0.151	0.330	0.014	0.092	0.471
Quarterly skip	0.039	0.150	0.316	0.006	0.073	0.454
Annual skip	0.056	0.121	0.287	0.011	0.057	0.423
Quarterly avg	0.067	0.218	0.568	0.012	0.074	0.437
Annual avg	0.077	0.124	0.292	0.011	0.065	0.492
$\rho = 0.98$						
Monthly	0.043	0.108	0.265	0.011	0.073	0.476
Quarterly skip	0.042	0.093	0.265	0.010	0.066	0.443
Annual skip	0.061	0.107	0.227	0.016	0.058	0.417
Quarterly avg	0.075	0.174	0.551	0.020	0.059	0.477
Annual avg	0.070	0.119	0.246	0.007	0.055	0.478

results from the cointegration analysis of the system (4.13) when $u_{2,t}$ is based on Blanchard. Again, we obtain similar results from λ_{max} and the $Trace$ statistics, therefore we only report the results from the λ_{max}. First of all, we notice that for every sample the Johansen test yields stronger results than the Engle-Granger. In fact, the probability to detect a cointegration relationship using the Johansen procedure is always higher than 90% except for the sample of annual data for 33 years. This complicates our analysis and we are not able to decide which feature affects the ability of the Johansen test more. However we can conclude that using averaged quarterly over annual observations for 33 years yields greater power gains than using annual observations for 66 years.

Although the Engle-Granger lacks of power we can pronounce over the importance of the frequency of observations. It is actually preferable to use monthly data of 33 years than quarterly, data acquired with the method of skip sampling, for 99 years. We can also conclude that if we need to lower the frequency of observations we better use the averaging with non-overlapping method.

We receive similar results when Evans' model is used to generate $u_{2,t}$. We still can not safely decide which of the two characteristics affects the power of the Johansen test more. As in the previous case, the ability of the Engle-Granger test depends more on the frequency of observations. The only exceptions concerns the use of

Table 5.11: Cointegration analysis of system 4.13

	Johansen			*Engle − Granger*		
	33 years	66 years	99 years	33 years	66 years	99 years
Model based on Blanchard (1979)						
Monthly	0.943	0.950	0.958	0.222	0.350	0.448
Quarterly skip	0.941	0.939	0.951	0.099	0.150	0.198
Annual skip	0.677	0.925	0.952	0.006	0.031	0.030
Quarterly avg	0.935	0.938	0.959	0.150	0.252	0.298
Annual avg	0.639	0.901	0.940	0.012	0.106	0.119
Model based on Evans (1991)						
Monthly	0.948	0.953	0.951	0.198	0.580	0.702
Quarterly skip	0.940	0.955	0.938	0.163	0.278	0.336
Annual skip	0.657	0.921	0.934	0.006	0.034	0.029
Quarterly avg	0.938	0.949	0.940	0.184	0.332	0.397
Annual avg	0.559	0.911	0.921	0.013	0.091	0.117

annual data for 66 years over monthly data for 33 years. We need around 41 years of monthly data to yield similar power gains as 66 years of quarterly data.

Chapter 6

Empirical applications

6.1 Coffee prices

To illustrate our findings from the simulations, we investigate the relationship between coffee prices. We use time series information on the four composite "indicator prices" constructed by the International Coffee Organization (ICO). These are spot coffee prices of Colombian milds, other milds, Brazilian unwashed arabica, and robusta. A similar analysis has been conducted by Vogelvang (1992) and revealed the existence of two cointegrating vectors.

The data set consists of monthly data over the 1965-2018 period and was kindly provided to us by professor Jesus Otero. Using the monthly data we compute quarterly and annual versions by skip sampling and averaging techniques. We then employ the Johansen test to test for cointegration for three sampling periods: 1965-2018, 1983-2018, 2001-2018.

Table 6.1: Cointegration analysis of coffee prices using Johansen test

	Annual		Quarterly		Monthly	
r	λ_{max}	Trace	λ_{max}	Trace	λ_{max}	Trace
Period: 1965-2018						
0	36.01***	64.69***	39.98***	76.64***	42.71***	89.80***
1	14.51	28.68	23.78**	36.66***	33.09***	47.09***
2	9.07	14.16	8.05	12.88	8.52	14.00
Period: 1983-2018						
0	28.86**	51.28**	28.26**	53.34**	30.32**	66.91***
1	13.12	22.42	14.05	29.36*	20.98*	36.60**
2	8.00	9.30	10.44	14.93	10.28	15.62
Period: 2001-2018						
0	18.08	28.93	15.83	37.27	24.76	58.61

The values denote the test statistics for r_0, r_1 and r_2. ***, ** and * denote significance at the 1, 5 and 10% significance levels, based on the critical values tabulated by MacKinnon et al. (1999)

Preliminary analysis revealed that all four variables are integrated of first order. Table 6.1 reports the results from Johansen test for the averaging method. We use r to denote the number of cointegrating vectors under the null hypothesis. We rely on the Schwartz criterion for the VAR order. There is no evidence of a cointegrating relationship for the shortest period of time. Over the period 1983-2018 both the λ_{max} and the $Trace$ tests detect two vectors when we use monthly data and one vector we use annual data. Applied on quarterly data the λ_{max} tests identifies only one vector while the $Trace$ test identifies two. Using the longest sample period of 53 years we are able to detect two vectors using monthly and quarterly observations and one cointegrating vector using annual data. We obtain similar results when we use the averaging with non-overlapping observations technique.

6.2 House prices and GDP

In the previous example we investigated the long-run relationship of four stationary variables. We now employ the Johansen method, to test for cointegration between variables that exhibit explosive behaviour. Specifically, we examine the cointegration relationship of houses prices in the U.S.A. and the GDP. Knoll et al. (2017) has studied the evolution of house prices in the long-run. Based on extensive data collection they proved that in most industrial economies houses prices rose sharply in the last decades. We obtain annual data from Òscar Jordà et al. (2016) that cover the 1918-2016 period. For the quarterly and monthly data we presume that the values of observations follow a linear progress. We test 3 samples as in the Monte Carlo simulations.

Preliminary analysis revealed the existence of a bubble in the house prices during the period 1944-2016. For the GDP there is evidence for two bubbles. One during the years 1942-1945 and a longer one during the period 1965-2016. For the period of 33 years the test rejects cointegration for annual and quarterly data. For monthly data we accept a cointegration relationship only for a 5% significance level. The test supports the presence of a cointegrating vector for the two longest sample periods.

Table 6.2: Cointegration analysis of house prices and GDP

Type of data	1984-2016	1951-2016	1918-2016
Monthly:			
λ_{max} test	15.64**	27.39***	43.91***
Trace test	18.20**	39.05***	48.87***
Quarterly:			
λ_{max} test	9.16	27.72***	43.89***
Trace test	12.85	33.17***	48.76***
Annual:			
λ_{max} test	11.54	35.51***	56.15***
Trace test	15.50	39.26***	58.94***

The values present the test statistic for the null hypothesis of no cointegrating vectors. *** and ** denote the significance at the 1 and 5% significance levels, based on the critical values tabulated by MacKinnon et al. (1999)

Chapter 7

Conclusions

Practitioners often have to decide whether to use monthly, quarterly or annual data when testing for unit roots and cointegration. In this research, we contact a Monte Carlo simulation experiments to examine the effects of increasing the frequency of observation and the data span on the Engle-Granger, the Phillips-Ouliaris and the Johansen cointegration tests. We considered cases of systems containing two, three and four time series and subcases considering the number of existing cointegrating vectors. In all cases the results indicate that when interested in long-run equilibrium, we ought to rely on data collected over a long period of time, rather than on a large number of data collected over a relatively short period of time. This findings are also supported by an empirical example where we investigated the relationship between coffee prices.

In addition, we also undertake simulation experiments that ought to be interesting for applied empirical work, such as the cases where practitioners face the presence of explosive behaviour, non-linearities and structural breaks in the underlying variables of interest. We mostly focus on explosive behaviour. Particularly, we investigate the effect of temporal disaggregation on the SADF and GSADF tests and on the Johansen and the Engle-Granger cointegration tests when encounter variables which exhibit explosive behaviour. Contrary to the previous results, the two bubble tests perform better when a higher frequency of observations is used. Additionally, the cointegration tests depend more on the data frequency when the variables contain multiple of incidents of periodically collapsing bubbles. On the contrary, when only one mildly explosive bubble exists, the two tests depend more on the total sample length.

In our research we simulated "monthly" data and obtained the "quarterly" and "annual" observations using the methods of skip sampling and averaging. However one could begin with low frequency data and construct higher frequency data using the different methods, such as linear or quadratic interpolation. One could test whether it is safe to fill missing values using these methods and compare the most probably increased power of the tests to the power of the tests when actual high frequency data is used.

Finally one could complement the existing literature by considering additional assumptions regarding the data. For example one could examine the effects of changing the data span and the frequency of observations on stock data which exhibits characteristics such as volatility clustering and fat tails.

Bibliography

Bagnai, A. (2008). Structural changes, cointegration and the empirics of thirlwall's law. *Applied Economics*, 37:1–15.

Bent, N. (2000). The asymptotic distribution of likelihood ratio test statistics for cointegration in unstable vector autoregressive processes.

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

Blanchard, O. J. and Watson, M. W. (1982). Bubbles, rational expectations and financial markets. Working Paper 945, National Bureau of Economic Research.

Diba, B. and Grossman, H. (1988). Explosive rational bubbles in stock prices? *American Economic Review*, 78(3):520–30.

Dickey, D. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057–72.

Doornik, J. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys*, 12(5):573–93.

Engle, R. and Granger, C. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–76.

Evans, G. (1991). Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review*, 81(4):922–30.

Granger, C. (1983). Co-integiated variables and error-correcting models. *UCSD Discussion Paper*, pp. 83-13a.

Gürkaynak, R. (2005). Econometric tests of asset price bubbles: taking stock. Finance and Economics Discussion Series 2005-04, Board of Governors of the Federal Reserve System (US).

Hamilton, J. and Whiteman, C. (1985). The observable implications of self-fulfilling expectations. *Journal of Monetary Economics*, 16(3):353–373.

Homm, U. and Breitung, J. (2011). Testing for speculative bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Econometrics*, 10:198–231.

Hooker, M. A. (1993). Testing for cointegration: Power versus frequency of observation. *Economics Letters*, 41(4):359 – 362.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2):231 – 254.

Juselius, K. and Mladenovic, Z. (2002). High inflation, hyperinflation and explosive roots. the case of yugoslavia. *University of Copenhagen. Department of Economics, Discussion Papers.*

Kapetanios, G., Shin, Y., and Snell, A. (2003). Testing for a unit root in the nonlinear star framework. *Journal of Econometrics*, 112(2):359–379.

Knoll, K., Schularick, M., and Steger, T. (2017). No Price Like Home: Global House Prices, 1870-2012. *American Economic Review*, 107(2):331–353.

Koehler, A. and S. Murphree, E. (1988). A comparison of the akaike and schwarz criteria for selecting model order. *Applied Statistics*, 37:187.

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

Lahiri, K. and Mamingi, N. (1995). Testing for cointegration: Power versus frequency of observation — another view. *Economics Letters*, 49(2):121 – 124.

Mackinnon, J. (1990). Critical values for cointegration tests. *Long-Run Economic Relationships*.

MacKinnon, J., Haug, A., and Michelis, L. (1999). Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics*, 14(5):563–77.

Narayan, P. and Sharma, S. (2015). Does data frequency matter for the impact of forward premium on spot exchange rate? *International Review of Financial Analysis*, 39.

Nelson, C. and Plosser, C. (1982). Trends and random walks in macroeconmic time series: Some evidence and implications. *Journal of Monetary Economics*, 10(2):139–162.

Nielsen, B. (2010). Analysis of coexplosive processes. *Econometric Theory*, 26(3):882–915.

Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, 54(3):461–72.

Otero, J. and Smith, J. (2000). Testing for cointegration: power versus frequency of observation — further monte carlo results. *Economics Letters*, 67(1):5 – 9.

Pfaff, B. (2006). *Analysis of integrated and cointegrated time series with R*. Springer.

Phillips, P. C. B. and Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58(1):165–193.

Phillips, P. C. B., Shi, S., and Yu, J. (2013). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. *International Economic Review*, 56(4):1043–1078.

Phillips, P. C. B., Wu, Y., and Yu, J. (2011). Explosive behavior in the 1990s nasdaq: When did exuberance escalate asset values? *International Economic Review*, 52(1):201–226.

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

Shiller, R. J. (1980). Do stock prices move too much to be justified by subsequent changes in dividends? Working Paper 456, National Bureau of Economic Research.

Shiller, R. J. and Perron, P. (1985). Testing the random walk hypothesis: Power versus frequency of observation. *Economics Letters*, 18(4):381 – 386.

Vogelvang, E. (1992). Hypotheses testing concerning relationship between spot prices of various types of coffee. *Journal of Applied Econometrics*, 7(2):191–201.

West, K. D. (1988). Bubbles, Fads, and Stock Price Volatility Tests: A Partial Evaluation. NBER Working Papers 2574, National Bureau of Economic Research, Inc.

Zhou, S. (2001). The power of cointegration tests versus data frequency and time spans. *Southern Economic Journal*, 67(4):906–921.

Zivot, E. and Andrews, D. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business Economic Statistics*, 10(3):251–70.

Òscar Jordà, Schularick, M., and Taylor, A. M. (2016). Macrofinancial History and the New Business Cycle Facts. In *NBER Macroeconomics Annual 2016, Volume 31*, NBER Chapters, pages 213–263. National Bureau of Economic Research, Inc.

Appendix A

Alternative cointegration cases including explosive series

During our research we conducted additional simulations which we briefly discuss here. In the first alternative case we implant the bubble to series x_t instead of y_t. The probability of identifying a cointegrating vector is lower but the general behaviour of the test is not affected.

Table A.1: Power of Johansen test when bubble appears on x

$\rho = 0.85$	33 years	66 years	99 years
Monthly	0.038	0.134	0.284
Quarterly skip	0.034	0.120	0.275
Annual skip	0.043	0.112	0.242
Quarterly avg	0.049	0.189	0.571
Annual avg	0.053	0.125	0.242
$\rho = 0.92$	33 years	66 years	99 years
Monthly	0.036	0.110	0.268
Quarterly skip	0.036	0.110	0.271
Annual skip	0.043	0.102	0.224
Quarterly avg	0.050	0.182	0.553
Annual avg	0.052	0.110	0.222
$\rho = 0.95$	33 years	66 years	99 years
Monthly	0.030	0.092	0.253
Quarterly skip	0.031	0.095	0.260
Annual skip	0.045	0.104	0.225
Quarterly avg	0.046	0.173	0.543
Annual avg	0.062	0.109	0.229
$\rho = 0.98$	33 years	66 years	99 years
Monthly	0.032	0.078	0.223
Quarterly skip	0.032	0.075	0.228
Annual skip	0.046	0.081	0.200
Quarterly avg	0.049	0.155	0.531
Annual avg	0.059	0.104	0.222

In the second case both time series contain a bubble during the same period. We observe a lower probability compared to the initial results but again the significance of a longer data length is more important than the number of observations. We obtain similar results when both time series contain a bubble at the same period.

Table A.2: Bubbles coincide

$\rho = 0.85$	33 years	66 years	99 years
Monthly	0.056	0.119	0.328
Quarterly skip	0.044	0.118	0.304
Annual skip	0.056	0.110	0.262
Quarterly avg	0.067	0.194	0.573
Annual avg	0.070	0.119	0.278
$\rho = 0.92$	33 years	66 years	99 years
Monthly	0.039	0.107	0.315
Quarterly skip	0.041	0.112	0.306
Annual skip	0.053	0.116	0.253
Quarterly avg	0.065	0.192	0.564
Annual avg	0.067	0.119	0.272
$\rho = 0.95$	33 years	66 years	99 years
Monthly	0.035	0.107	0.299
Quarterly skip	0.038	0.111	0.297
Annual skip	0.047	0.106	0.245
Quarterly avg	0.061	0.183	0.553
Annual avg	0.060	0.119	0.264
$\rho = 0.98$	33 years	66 years	99 years
Monthly	0.033	0.084	0.253
Quarterly skip	0.029	0.090	0.243
Annual skip	0.051	0.091	0.217
Quarterly avg	0.052	0.169	0.531
Annual avg	0.059	0.099	0.232

The last case we examined concerns two explosive time series with bubbles appearing at different periods. In series y_t the explosive process takes time during years 10-25 and in x_t during the years 40-55. The results we obtained using a 33 years sample are similar to those obtained when only y_t contained a bubble. Using a 66 years sample yields power loss up to 50% compared to the initial simulations. Using the longest sample period yields power gains but the probability of identifying a cointegrating vector remains low.

Table A.3: Bubbles appear at different periods

	33 years	66 years	99 years
Monthly	0.056	0.054	0.179
Quarterly skip	0.044	0.062	0.167
Annual skip	0.056	0.059	0.145
Quarterly avg	0.067	0.132	0.531
Annual avg	0.070	0.071	0.166
$\rho = 0.92$	33 years	66 years	99 years
Monthly	0.039	0.057	0.174
Quarterly skip	0.041	0.059	0.166
Annual skip	0.053	0.061	0.144
Quarterly avg	0.065	0.134	0.517
Annual avg	0.067	0.075	0.160
$\rho = 0.95$	33 years	66 years	99 years
Monthly	0.035	0.047	0.158
Quarterly skip	0.045	0.048	0.151
Annual skip	0.053	0.061	0.148
Quarterly avg	0.058	0.131	0.527
Annual avg	0.064	0.063	0.162
$\rho = 0.98$	33 years	66 years	99 years
Monthly	0.039	0.046	0.174
Quarterly skip	0.043	0.052	0.158
Annual skip	0.054	0.065	0.143
Quarterly avg	0.064	0.142	0.532
Annual avg	0.064	0.074	0.166