Essays on Specification Testing in Time Series with Applications to Statistical Arbitrage

Oron Daihes [1]

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

July 2012

[1]The author can be contacted at Oron@Daihes.com

In memory of my grandmother Efrat Hodak. What a woman of class, power, soul and dignity. How much I love her.

Contents

1 Introduction **5**

2 Testing for unit roots in the presence of polynomial trends **12**

2.1 Introduction 14

2.2 The model 18

2.3 Asymptotic Behavior Under Cubic Trends 19

2.4 Union of Rejection Strategy 22

2.5 Finite Sample Simulations 24

2.6 Application to Commodity Prices 25

2.7 Statistical Arbitrage and Testing for Cointegration 28

2.8 Conclusion 32

3 The robustness of a panel stationarity test to nonstationary volatility **53**

3.1 Introduction 55

3.2 The model 57

3.2.1 The estimator 59

3.3 Monte Carlo simulations 61

3.4 Conclusion 66

4 Detecting multiple level breaks in the presence of non-stationary volatility 86

4.1 Introduction 88

4.2 A multiple level breaks model with non-stationary volatility . 89

4.3 The HLT procedure for testing for multiple level breaks 93

4.3.1 Estimation of the Long Run Variance of ε_t and u_t . . . 95

4.3.2 The HLT tests 98

4.4 Accounting for non-stationary volatility 101

4.4.1 Wild bootstrap procedure I 103

4.4.2 Wild bootstrap procedure II 105

4.4.3 Difficulty of bootstrap for the U test 106

4.4.4 The wild bootstrapped U test for level breaks 107

4.4.5 Problem with wild bootstrap procedure I 108

4.4.6 Simulation results 114

4.5 Conclusion 116

5 Statistical arbitrage with stochastic cointegration 128

5.1 Introduction 129

5.2 Stochastic Integration 134

5.2.1 Stochastic Cointegration 136

5.2.2 Hypothesis Tests and Test Statistics 138

5.3 Empirical Methodology and Results 143

5.4 Conclusion 152

ACKNOWLEDGMENTS

My first, and most earnest, acknowledgment must go to my advisor Prof. Steve Leybourne for his encouragement, advice, mentoring, and research support throughout my doctoral studies. In every sense, none of this work would have been possible without him.

I also owe a huge debt of gratitude to Prof. Oliver Morrissey for enlightening discussions we had towards the end of my Master's studies, encouraging me to further explore in great depth, some important theoretical aspects in econometrics as well as their applications to finance and economics.

A special acknowledgment must go to my wonderful parents. For always being there when I needed them most.

My final, and most heartfelt, acknowledgment must go to my wife Kristen. Kristen has worked diligently, and successfully, for four years to show me life outside of the academia. Her support, encouragement, and companionship have turned my journey through my doctoral studies into a pleasure. For all that, and for being everything I am not, she has my everlasting love.

Chapter 1

Introduction

Quantitative methods of speculation have long been an interest to major market participants. "Quants", those who exploit quantitative methods, usually use sophisticated mathematical and statistical techniques to develop high-tech trading programs, executable through automated trading systems and achieved by disciplined, consistent rules. Determining the correct specification of a time series is important both in economics and as a practical tool for a trader. The goal of the economist is to study certain aspects of the economy and provide accurate predictions about the effect of structural changes in the economy, unexpected shocks to the fundamentals of the economy or public policy proposals. The goal of an arbitrageur is to generate financial gain by arbitraging significant deviations from the modeled phenomena. In both cases an accurate specification of time series data is needed.

Test statistics that are usually developed for basic models do not apply to real data. It is often the case that after analyzing a particular dataset in a descriptive way it becomes apparent that the basic model fails to capture certain aspects of the data. Thus the test statistic needs to be analyzed under certain well-defined departures from the basic model. Sometimes, however, the tests remain relatively robust to a more general model. In Chapter 3, for example, I find that tests for stationarity in panel data that assume a stationary volatility process are in the least favorable case only modestly oversized in instances where the volatility process is nonstationary. Those tests in fact retain power quite well. On the other hand, tests that are developed for a simple model often exhibit severe size distortions when the data departs from this simple model. When this is the case these statistics can be extended to a broader more relevant class of models in several ways.

For example, tests of nonstationarity rely on assumptions about a time trend, if one is present, and assumptions about the dynamics of the underlying driving process. In Chapter 2 I find that failure to account for the time trend indeed leads to a test with zero size and zero power. Correcting the statistic entails properly accounting for the time trend. A recent line of literature is concerned with the properties of these tests in the presence of uncertainty about the existence and form of the trend (Harvey, Leybourne and Taylor 2009b, Harvey, Leybourne and Taylor 2011). If it is "underspecified" the test has (close to) zero size and power. If it is "overspecified" the test has correct size but inferior power. These authors propose a union of rejections testing strategy to handle the trend uncertainty. Chapter 2 discusses this work and considers time series with a cubic trend. I show that the union of rejections strategy can be extended to this more general model.

Sometimes the uncertainty over the correct specification can be solved in yet a different way. When the correct null distribution depends on unknown parameters a resampling technique can often be used to approximate the distribution. When testing for structural level breaks in a time series the correct specification of the volatility process affects the null distribution of the statistic. Specifying a parametric model for the volatility process is beneficial if one has exact knowledge of the data generating process (dgp). In Chapter 3, I instead implement a nonparametric, bootstrapping procedure that accounts for the volatility process without a parametric structure.

These specification tests have real practical importance. There is evidence that many macroeconomic time series are characterized by either permanent or transitory shocks fluctuating around a long-run mean that exhibits sudden breaks (Stock and Watson 1996). Moreover, Busetti and Taylor (2003) and Sensier and Dijk (2004) find that series with level breaks tend to also exhibit breaks in volatility. Indeed a large body of recent work has shown that the unconditional volatility of the processes underlying many macroeconomic time series declined over the last quarter of the last century (see, e.g., the literature review in Cavaliere and Taylor, 2008). Thus the importance of simultaneously accounting for uncertainty over the order of integration and the possibility of breaks in volatility when testing for level breaks - as the procedure I develop in Chapter 4 does - cannot be overstated.

Accounting for trend uncertainty when testing for a unit root is also empirically important. Many macroeconomic times series exhibit a combination of a nonstochastic time trend and nonstationary stochastic behavior. In Chapter

2 I analyze one hundred years of data on relative commodity prices. These prices evidently exhibit a nonstochastic trend, often downward. In addition, Harvey et al. (2011) show that many of the commodity price series can be described as stationary around this nonstochastic trend. Determining the order of integration of these series is an important task for economic forecasters who wish to predict the affect of economic shocks or policy changes. Our analysis in Chapter 2 demonstrates that accounting for uncertainty over the form of the trend is empirically relevant as it leads to different conclusions as to the order of integration of these commodity prices. For example, I find that the price of sugar is stationary around a cubic trend whereas Harvey et al. (2011) concluded that it was nonstationary.

Another practical application of specification testing arises in quantitative methods of speculation by actual market participants. A popular short term speculation strategy that belongs at the arsenal of statistical arbitrage tools and is currently used by hedge funds as well as investment banks, known as "pairs trading". The underlying concept is simply to identify two stocks, or other traded assets, whose prices have moved together historically. When the spread between them widens, a profitable action would be to short the winner and long the loser. If the co-movement of the two assets is correctly specified then the prices will converge and the arbitrageur will profit.

Asset pricing can be viewed in absolute and relative terms. Absolute pricing values securities based on fundamentals such as discounted future cash flows. This method is often notoriously difficult and results in wide margin for error. Relative pricing means that two securities that are close substitutes for each other should sell for a similar price. Relative pricing is only slightly easier as it does not say what the price will be; it does, however, enable to infer the price of one asset given the price of the other. The Law of One Price (LOOP) is a term associated with relative pricing, described by Install (1987): "two investments with the same payoff in every state of the nature must have the *same* current value". Differently stated, two securities with the exactly same prices in all states of the world should sell for exactly same price. Chen and Knez (1995) extend this argument and posit that *closely integrated* markets should assign to similar payoffs prices that are close. This weaker condition implies that two securities with similar, but not necessarily identical payoffs across states should have *similar* prices. This proposition allows the examination of near-efficient economies, or in Chen and Knezs terminology, near integrated markets. In markets which are efficient, risk-adjusted returns from pure arbitrage strategies should not be positive. However, the fact that speculators are trading securities that are close economic substitutes for each other may validate the existence of short term arbitrage opportunities. Specification testing plays an important role here by identifying these assets that co-move, as well as identifying when the spread has widened sufficiently to enable arbitrage. Controlling the size of the relevant tests is a risk-management tool; it safeguards against too often falsely detecting opportunities for arbitrage. Increasing the power of the relevant tests increases the arbitrage opportunities detected and hence increases profitability of a strategy.

Detecting assets whose prices have moved together historically depends crucially on how this "co-movement" is defined. The simple correlation between the two assets measures co-movements in returns. However, it measures the degree of linear association over time and does not carry information on the long-term relationship that may exist between the variables. The cointegration framework of Engle and Granger (1987) is a more useful concept of co-movement in returns. Cointegration measures *long-term* co-movements in prices, which prevails even through periods in which static correlation is low.

Two time series are said to be cointegrated if they are each $I(1)$ but some linear combination of the two is $I(0)$. Hence statistical tests for cointegration relationships can be derived from unit root and stationarity tests. In Chapter 2 I apply the unit root tests that account for uncertainty over the trend to a test for cointegration between two assets when one or possibly both assets exhibit a linear, quadratic or even cubic time trends. Not accounting for the possibility of time trends could lead to the failure to detect cointegrating relationships and hence limit the profitability of pairs trading.

In Chapter 5 I discuss another important issue related to testing for cointegrated assets. In some cases the cointegration framework of Engle and Granger (1987) fails to detect cointegrating relationships when the assets appear to co-move historically. This can happen for instance if the spread between them "moves too much" (see Campbell and Shiller (1987)). Harris, McCabe and Leybourne (2002) and McCabe, Leybourne and Harris (2006) developed a framework of *stochastic cointegration* to account for this possibility. Two time series are said to be stochastically cointegrated if some linear combination of the two series is *stochastically trendless*. This term is defined precisely in Chapter 5. This framework allows for nonstationary heteroskedasticity in the spread so that shocks may have a permanent effect on the variance but always have only a transitory effect on the level of the spread. In Chapter 5 I show that this is an effective tool that consistently estimates long-run states of equilibrium and repeatedly detects such relationships in the U.S. equity market. A test for stationarity of the spread between two assets that does not allow for nonstationary heteroskedasticity rejects the null hypothesis of co-movement too often. Accounting for the possibly for nonstationary heteroskedasticity properly controls the size of this test. This reinforces the central theme of this thesis that careful specification testing of economic or financial time series is valuable and important.

In this thesis I study several specification tests and analyze ways of extending them in order to enhance their practical utility both for economists and traders. The second chapter studies unit root tests in the presence of uncertainty about the non-stochastic time trend in the data. The third and fourth chapters consider the role of nonstationary volatility in specification tests. Chapter 3 shows that a test of stationarity in a panel exhibits a surprising robustness in the face of time-varying variances. In Chapter 4, after finding that this is not true for a test for level breaks, I propose a solution based on resampling techniques. Chapter 5 is empirical. I demonstrate that the stochastic cointegration framework of Harris et al. (2002) and McCabe et al. (2006) is better able to detect cointegrating relationships in the U.S. equity market.

Chapter 2

Testing for unit roots in the presence of polynomial trends

The specification of the deterministic trend component of a time series is an important step in testing for the presence of a unit root. Recent work by Ayat and Burridge (2000) emphasizes the importance of allowing nonlinear trends. Harvey et al. (2009b) and Harvey et al. (2011), as well as Ayat and Burridge (2000), describe problems associated with testing the null hypothesis of a unit root when there is uncertainty about the trend component. Harvey et al. (2009b) and Harvey et al. (2011) described testing strategies based on a union of rejections approach that can circumvent the low power associated with overfitting the model for the trend and demonstrate this approach by allowing for no trend, a linear trend or a quadratic trend. Their analysis suggests that this approach can be extended to allow more flexible specifications. In this paper we develop a unit root testing strategy that allows for nonlinearity in the form of cubic trends in the time series. We develop the asymptotic theory of the test and provide simulations that show substantial improvement of the union test over a test based on a cubic trend estimate, both asymptotically and in finite samples. We show that this testing strategy is important empirically by applying it to two separate data analyses. First we use the methodology to test for unit roots in a series of commodity price data, analyzed in Harvey et al. (2011). Second, we test for cointegration between every pair combinations of the 30 stocks that constitute the Dow Jones Industrial Average index, and find that the procedure detects far more cointegrating relationships, when compared with the linear or quadratic specifications. Thus, extending the possibilities to statistically arbitrage stocks that are sufficiently away from their long-run estimated equilibria.

2.1 Introduction

Testing for the presence of a unit root is an important issue in time series analysis. The literature on unit root tests has been very active since the influential Dickey and Fuller (1979) paper derived a unit root test based on the t-test for the null hypothesis that the regression estimate of the AR coefficient is equal to 1. One problem with the Dickey-Fuller test and extensions to it that have been discussed in the literature, such as the Elliott, Rothenberg and Stock (1996) test that was shown to be efficient under certain conditions, is that they are sensitive to the assumed deterministic component of the time series. For example, Perron (1989) studied the sensitivity of unit root tests to breaks in the deterministic trend. Overfitting the deterministic component generally leads to low power of the test for a unit root. On the other hand, underfitting the deterministic component will also generally affect the power of the test.

Thus the specification of the form of the trend component is very important. There are two separate issues at stake. One is an issue of interpretation and the other is an issue of statistical power. If a very flexible deterministic trend is allowed then it is not clear what evidence of stationarity really means. In the extreme case, the observed time series can be entirely attributed to a nonparametric deterministic trend. Since all that is left after removing such a trend is $y_t = 0$, the series is trivially stationary. A separate but related issue is that allowing a highly parameterized model for the deterministic trend introduces a lot more statistical noise. In the presence of a lot of noise any statistical test for a unit root is likely to have very low power. We only deal with the issue of statistical power in this paper, while keeping in mind that a more flexible deterministic trend model may alter the interpretation of the results of the tests.

Many authors have considered flexible models for the deterministic trend. Ouliaris, Phillips and Park (1989) derive a unit root test that allows for a polynomial trend of given degree. Though they do not discuss the asymptotic (local) power of their test it is clear from their applications to real data that the test has low power. Bierens (1997) estimates the trend nonparametrically via Chebyschev polynomials. The results of the proposed tests when applied to several macroeconomic indicators are difficult to interpret because the trends swallows nearly the whole series. But in general he does not find low rejection rates. Becker, Enders and Lee (2006) use a flexible Fourier approximation. Other recent contributions to this literature include Pippenger and Goering (1993), Balke and Fomby (1997), and Kapetanios, Shin and Snell (2003).

In a pair of papers Harvey, Leybourne and Taylor study unit roots tests when the deterministic component may be linear or quadratic (Harvey et al. 2009b, Harvey et al. 2011). Instead of allowing a very flexible specification of the trend component they focus on uncertainty between two parsimonious nested models for the trend. In some sense it seems that their approach simultaneously solves the problem of interpretation (discussed, for example, in Phillips (1998)) and low statistical power. The interpretation of their test is clear: does the time series have a unit root or is it stationary around a trend that is either linear or locally quadratic? It is clear that there is still substantial variation around the estimated trend in their commodity price application, as opposed to Bieren's estimate trends for various time series of price indices. However, the focus on two nested models in some way seems to exacerbate the problem of low statistical power. Harvey et al. (2009b) and Harvey et al. (2011) then solve this problem in a clever way by using a strategy that rejects when at least one of two different statistics suggests rejecting the null. Thus the important result that can be learned from these papers is that a limited amount of flexibility in the specification of the trend can be allowed without altering the interpretation and without sacrificing substantial power. Thus allowing a more complex model for the trend does not risk overfitting because their strategy also rejects if the less complex trend model would suggest rejection. This strategy works in part because if the more complex model were indeed correctly specified then the less complex model would nearly never reject the null. In Harvey et al. (2009b) the two models are no trend versus a linear trend; in Harvey et al. (2011) they are linear trend versus quadratic trend.

In this paper we extend this by allowing a cubic trend. This is done in part to show that the logic of these two papers can be extended to more flexible specifications of the deterministic trend. The question we leave on the table then is to what extent such an approach can be extended to allow a polynomial trend of unknown degree and how such an approach that carefully guards against loss of power would compare to the tests proposed by Ouliaris et al. (1989) and Bierens (1997). We first derive the asymptotic distribution of the Elliott et al. (1996) unit root test that allows for a linear trend and the Harvey et al. (2011) unit root test that allows for a quadratic trend under the assumption of a local cubic trend. As anticipated we find that probability of rejection under these tests quickly collapses to zero as the cubic component of the trend is increased. We next derive asymptotic properties of a new test statistic, $DF - QD^{cb}$, that differs from the others by allowing a third degree polynomial in the QD detrending step. As expected, this test is asymptotically invariant to the cubic component of the trend. We plot the power function and show the power that is sacrificed by using this test when the cubic trend component is very small or nonexistent. Finally, in the spirit of Harvey et al. (2009b) and Harvey et al. (2011) we propose a union of rejections strategy that takes advantage of the high power of the proper test under different conditions. We find that the size and power properties of these tests are not substantially different in finite samples.

The rest of the paper is organized as follows. In Section 2 we describe our time series model that allows for cubic time trends and define the Dickey-Fuller test statistics we use. In Section 3 we derive the asymptotic behavior of the various tests under a cubic trend. In Section 4 we lay out the union of rejections strategy. In Section 5 we report simulation results for the finite sample properties of the tests. Finally, Sections 6 and 7 provide applications to unit root tests on commodity price data and cointegration tests on stock price data, respectively. Section 8 concludes.

2.2 The model

Suppose we observe a time series of length T, $\{y_t\}$. The model, or data generating process, we assume is

$$y_t = \mu + \beta t + \gamma t^2 + \delta t^3 + u_t, t = 1, \ldots, T \tag{2.1}$$

$$u_t = \rho_T u_{t-1} + \varepsilon_t, t = 2, \ldots, T \tag{2.2}$$

where $\rho_T = 1 - cT^{-1}$, for $0 \leq c < \infty$, $T^{-1/2}u_1 \rightarrow_p 0$ and $\{\varepsilon_t\}$ is a stable and invertible linear process. Note that $c = 0$ corresponds to a unit root process and $c > 0$ corresponds to the local alternative. That is, we assume

Assumption 2.1. *The stochastic process* $\{\varepsilon_t\}$ *satisfies*

$$\varepsilon_t = C(L)e_t$$
$$C(L) := \sum_{i=0}^{\infty} C_i L^i, C_0 = 1$$

where $C(z) \neq 0$ *for all* z *such that* $|z| \leq 1$, $\sum_{i=1}^{\infty} i|C_i| < \infty$ *and* $\{e_t\}$ *is a martingale difference sequence with conditional variance* σ^2 *and* $sup_t E(e_t^4) < \infty$.

We would like to test the null hypothesis $H_0 : \rho = 1$ against the local alternative $H_1 : \rho < 1$, and consider the power under local alternatives where $\rho = \rho_T := 1 - c/T$ for $c > 0$.

Three different tests can be defined, depending on whether a linear, quadratic or cubic trend is assumed. Define $\bar{\rho}_{\bar{c}} := 1 - \bar{c}/T$. Let $z_t^{\tau} = (1, t)'$, $z_t^q = (1, t, t^2)'$

and $z_t^{cb} = (1, t, t^2, t^3)'$. Finally, let $\mathbf{y}_{\bar{c}} := (y_1, y_2 - \bar{\rho}_{\bar{c}} y_1, \ldots, y_T - \bar{\rho}_{\bar{c}} y_{T-1})'$ and $\mathbf{Z}_{\bar{c}}^i := (z_1^i, z_2^i - \bar{\rho}_{\bar{c}} z_1^i, \ldots, z_T^i - \bar{\rho}_{\bar{c}} z_{T-1}^i)'$ for $i = \tau, q, cb$. Then define $\tilde{u}_t^i := y_t - z_t^{i\prime} \tilde{\theta}^i$ where $\tilde{\theta}^i$ is the coefficient from the QD trend regression of $\mathbf{y}_{\bar{c}_i}$ on $\mathbf{Z}_{\bar{c}_i}^i$. Then $DF - QD^i$ is equal to the t-statistic for $\rho = 1$ in the regression

$$\tilde{u}_t^i = \rho \tilde{u}_{t-1} + \sum_{j=1}^{p} \phi_j \Delta \tilde{u}_{t-j}^i + e_t$$

Clearly we have defined $DF - QD^\tau$ to coincide with the Dickey and Fuller (1979)-type test proposed by Elliot et al (1996) and $DF - QD^q$ to coincide with the test recommended by Ayat and Burridge (2000). As far as we know $DF - QD^{cb}$ has not been used in the literature. Note that we have allowed the quasi-differencing parameter $\bar{c}$ to be different for the three tests. The "optimal" values have been found to be $\bar{c}_\tau = 13.5$, $\bar{c}_q = 18.5$ and $\bar{c}_{cb} = 23$[1]

2.3 Asymptotic Behavior Under Cubic Trends

In this section we give the asymptotic distribution of all three test statistics under the assumption of a cubic trend. In order to prevent the cubic and quadratic terms in the trend from dominating the asymptotic behavior we need to assume that these coefficients are decaying toward zero as a function of the sample size at the appropriate rate. The relevant Pitman drifts are $\gamma_T = \kappa_1 T^{-3/2}$ and $\delta_T = \kappa_2 T^{-5/2}$. Under these sequences of parameter values the asymptotic distributions of the three test statistics introduced in the previous

[1]The first two have been reported in Elliott et al. (1996) and Harvey et al. (2011). We find via simulations that the Gaussian power envelope in the cubic trend case is at 0.50 for $c = 23$.

section can be characterized in terms of κ_1 and κ_2. The following lemma gives the asymptotic distributions.

Lemma 1. *Suppose $\{y_t\}_{t=1}^T$ is generated according to equations (2.1)-(2.2) and that ρ, γ and δ are indexed by T: $\rho_T = 1 - c/T$, $\gamma_T = \kappa_1 T^{-3/2}$, and $\delta_T = \kappa_2 T^{-5/2}$. Then under this sequence of distributions the test statistics have the following limits*

$$DF - QD^i \xrightarrow{d} \frac{J_c^i(1)^2 - 1}{2\sqrt{\int_0^1 J_c^{i,\bar{c}_i}(r)^2 dr}} := \tau_i$$

where

$$\begin{aligned}
J_c^{\tau,\bar{c}_\tau} &= W_c(r) - r\pi_{1,\bar{c}_\tau}^{-1} M_{1,\bar{c}_\tau} + \kappa_1^*(r^2 - r\pi_{1,\bar{c}_\tau}^{-1}\pi_{2,\bar{c}_\tau}) + \kappa_2^*(r^3 - r\pi_{1,\bar{c}_\tau}^{-1}\pi_{4,\bar{c}_\tau}) \\
J_c^{q,\bar{c}_q} &= W_c(r) - rd_{\bar{c}_q}^{-1}(\pi_{3,\bar{c}_q} M_{1,\bar{c}_q} - \pi_{2,\bar{c}_q} M_{2,\bar{c}_q}) - r^2 d_{\bar{c}_q}^{-1}(\pi_{1,\bar{c}_q} M_{2,\bar{c}_q} - \pi_{2,\bar{c}_q} M_{1,\bar{c}_q}) \\
&\quad + \kappa_2^*(r^3 - rd_{\bar{c}_q}^{-1}(\pi_{3,\bar{c}_q}\pi_{4,\bar{c}_q} - \pi_{2,\bar{c}_q}\pi_{5,\bar{c}_q}) - r^2 d_{\bar{c}_q}^{-1}(\pi_{1,\bar{c}_q}\pi_{5,\bar{c}_q} - \pi_{2,\bar{c}_q}\pi_{4,\bar{c}_q}) \\
J_c^{cb,\bar{c}_{cb}} &= W_c(r) - r(\pi^{(11)} M_{1,\bar{c}_{cb}} + \pi^{(12)} M_{2,\bar{c}_{cb}} + \pi^{(13)} M_{3,\bar{c}_{cb}}) \\
&\quad - r^2(\pi^{(21)} M_{1,\bar{c}_{cb}} + \pi^{(22)} M_{2,\bar{c}_{cb}} + \pi^{(23)} M_{3,\bar{c}_{cb}}) - r^3(\pi^{(31)} M_{1,\bar{c}_{cb}} + \pi^{(32)} M_{2,\bar{c}_{cb}} + \pi^{(33)} M_{3,\bar{c}_{cb}})
\end{aligned}$$

where $\kappa_s^* = \kappa_s/\sigma$ for $s = 1, 2$, $W_c(r) = \int_0^r \exp(-(r-s)c)dW(s)$ for standard Brownian motion $W(s)$ and

$$\begin{aligned}
M_{1,\bar{c}_i} &= (1 + \bar{c}_i)W_c(1) + \bar{c}_i^2 \int_0^1 sW_c(s)ds, \quad i = \tau, q, cb \\
M_{2,\bar{c}_i} &= (2 + \bar{c}_i)W_c(1) - 2\int_0^1 W_c(s)ds + \bar{c}_i^2 \int_0^1 s^2 W_c(s)ds, \quad i = q, cb \\
M_{3,\bar{c}_{cb}} &= (3 + \bar{c}_{cb})W_c(1) - 6\int_0^1 sW_c(s)ds + \bar{c}_{cb}^2 \int_0^1 s^3 W_c(s)ds
\end{aligned}$$

and for $i = \tau, q, cb$,

$$\begin{aligned}
\pi_{1,\bar{c}_i} &= 1 + \bar{c}_i + \bar{c}_i^2/3 \\
\pi_{2,\bar{c}_i} &= 1 + \bar{c}_i + \bar{c}_i^2/4 \\
\pi_{3,\bar{c}_i} &= 4/3 + \bar{c}_i + \bar{c}_i^2/5 \\
\pi_{4,\bar{c}_i} &= 1 + \bar{c}_i + \bar{c}_i^2/5 \\
\pi_{5,\bar{c}_i} &= 3/2 + \bar{c}_i + \bar{c}_i^2/6 \\
\pi_{6,\bar{c}_i} &= 9/5 + \bar{c}_i + \bar{c}_i^2/7 \\
d_{\bar{c}_i} &= \pi_{1,\bar{c}_i}\pi_{3,\bar{c}_i} - \pi_{2,\bar{c}_i}^2 \\
\Pi &= \begin{pmatrix} \pi_{1,\bar{c}_{cb}} & \pi_{2,\bar{c}_{cb}} & \pi_{4,\bar{c}_{cb}} \\ \pi_{2,\bar{c}_{cb}} & \pi_{3,\bar{c}_{cb}} & \pi_{5,\bar{c}_{cb}} \\ \pi_{4,\bar{c}_{cb}} & \pi_{5,\bar{c}_{cb}} & \pi_{6,\bar{c}_{cb}} \end{pmatrix} \\
\Pi^{-1} &= (\pi^{(ij)})
\end{aligned}$$

Note that the Pitman drift parameters κ_1^* and κ_2^* cannot be consistently estimated. The fact that τ_τ and τ_q depend on these parameters, while τ_{cb} does not, highlight the importance of using τ_{cb} when a cubic trend is known to be present. The other two tests will suffer from low power because the critical value cannot be indexed by κ_s^*.

Figure 1 shows the asymptotic power functions that we computed by direct simulation of the distributions in Lemma 1. The Brownian motions are approximated via standard normal random variables and the integrals of the relevant processes by partial sums over 500 steps. Also note that from Lemma

1 only the distribution of $DF-QD^{\tau}$ depends on κ_1^*. Since the interest here is in the relative behavior of $DF-QD^q$ and $DF-QD^{cb}$ we set $\kappa_1^* = \kappa_2^*$. We then plot the power functions for $\kappa_2^* \in \{0, 1, 2, 3, 4, 5\}$. Several main patterns can be seen. The reported results are for a nominal size of 5%. First, for $\kappa_2^* = 0$ the linear trend statistic is most powerful (because $\kappa_1^* = 0$) and the quadratic trend statistic is more powerful than the cubic. This dominance is maintained for small values of κ_2^*. And at $\kappa_2^* = 2$ $DF-QD^q$ and $DF-QD^{cb}$ have very similar power functions. For higher values of κ_2^* the power of $DF-QD^q$ collapses while the power for $DF-QD^{cb}$ is unchanged. Finally, in unreported simulations we saw that an increase in κ_1^* has a different affect on the power of $DF-QD^{\tau}$ than the same increase in κ_2^*. As one might expect, the power collapses faster with the local cubic trend coefficient κ_2^*.

Table 1. Critical Values

	0.1	0.05	0.01
Linear	*-2.56*	*-2.85*	*-3.41*
Quadratic	*-3.15*	*-3.43*	*-3.97*
Cubic	**-3.62**	**-3.89**	**-4.35**
HLT adj.	*1.069*	*1.058*	*1.043*
new adj.	**1.095**	**1.079**	**1.071**

2.4 Union of Rejection Strategy

From the results seen in Figure 1 it is apparent that if the researcher knew whether or not a cubic trend is present in the data the choice of test statistic is crucial but obvious. One possible solution is to pre-test for a cubic trend. However, we worry as do Harvey et al. that the low power of such a pre-test for moderate values of the cubic coefficient would result in low power of the unit root test.

Instead we follow Harvey et al. (2011) and propose a union of rejections strategy. Under this strategy the unit root hypothesis is rejected if it is rejected by *any one* of the three tests, $DF - QD^i$, $i = \tau, q, cb$. This is equivalent to rejecting if the following test statistic is less than cv_τ:

$$t_{UR} := \min\{DF - QD^\tau, \frac{cv_\tau}{cv_q} DF - QD^q, \frac{cv_\tau}{cv_c} DF - QD^{cb}\}$$

We denote the resulting test by $UR_{\tau,q,cb}$. In general $UR_{\tau,q,cb}$ will be oversized since $Pr(t_{UR} < cv_\tau) \geq Pr(DF - QD^\tau) = \alpha$ where the inequality is strict as long as the second term in the "min" is smaller with some probability. However, the critical value can be adjusted in order to correct the size. This adjustment is likely to sacrifice power because the adjusted test will always reject less often than the unadjusted test. The adjusted union of rejections strategy rejects the null if $t_{UR} < \psi cv_\tau$. We denote this test by $UR^{adj.}_{\tau,q,cb}$. Note that the size will be highest when $\kappa_1^* = \kappa_2^* = 0$. As a result, we can obtain the right value of ψ that will control the size by simulating the distribution under this assumption. We computed these values for $\alpha = 0.01$, 0.05 and 0.10 through a Monte Carlo simulation with 5000 iterations. The values of ψ were obtained via a grid search that gave size closest to the desired alpha. They are reported in the last row of Table 1 above (new adj.).

Figure 1 includes simulations of the power and size of these two tests as well. Also, the test $UR_{\tau,q}$ denotes the union of rejections test proposed by Harvey et al. (2009b) that is based on the $DF-QD^{\tau}$ and $DF-QD^{q}$ statistics alone, and $UR^{adj.}_{\tau,q}$ denotes the version of this test with the size correction[2].

As expected the unadjusted $UR_{\tau,q,cb}$ test maintains the highest possible power across all values of κ^*_1. This result follows directly from the definition of this procedure. The problem of course is that it will reject too often under the null. In fact the size of this test is 8.5% for a nominal 5% test. Note however, that the adjusted test still performs quite well across all values of κ^*_2.

Another important aspect of Figure 1 is the comparison between $UR^{adj.}_{\tau,q,cb}$ and the proposed procedure of Harvey et al. (2011), $UR^{adj.}_{\tau,q}$. The former certainly maintains much better power when $\kappa^*_2 \neq 0$. However, when $\kappa^*_2 = 0$ the (expected) power advantage of $UR^{adj.}_{\tau,q}$ is very small.

2.5 Finite Sample Simulations

Figure 2 shows the size and power of the tests in finite samples. We simulate a series from the model of equations (2.1)-(2.2) with $\varepsilon_t \sim iid \quad N(0,1)$, $T = 150$, $\rho = 1 - c/T$ with c ranging from 0 to 40 and with $\mu = \beta = \gamma = 0$ and $\delta = \kappa_2 T^{-5/2}$ for κ_2 ranging from 0 to 5. The number of lags used in the Dickey-Fuller regression are chosen according to the modified MAIC procedure suggested by Perron and Qu (2007).

Figure 2 depicts the finite sample power functions for $DF-QD^{\tau}$, $DF-QD^{q}$, $DF-QD^{cb}$, $UR_{\tau,q}$, $UR^{adj.}_{\tau,q}$, $UR_{\tau,q,cb}$, and $UR^{adj.}_{\tau,q,cb}$, approximated by 1000 Monte Carlo simulations. First note that the tests are not oversized in the

[2]As reported by Harvey et al. (2011) the appropriate adjustment for this test is $\kappa = 1.058$ for a 5% level test. For our test the appropriate adjustment is $\kappa = 1.079$ for a test of level 5%.

finite sample, except $UR_{\tau,q}$ and $UR_{\tau,q,cb}$ which are expected to be oversized as they are oversized asymptotically as well. For example, the size of $UR^{adj.}_{\tau,q,cb}$ ranges from 0.031 to 0.049.

Second, note that the tests maintain good power in the finite sample. As the asymptotic analysis suggests, the union of rejections strategy is able to capitalize on the superior power of $DF-QD^{\tau}$ when only a linear trend is present, the superior power of $DF-QD^{q}$ when only a quadratic trend is present, and the superior power of $DF-QD^{cb}$ when a cubic trend is present. The fact that the finite sample performance largely matches the asymptotic behavior suggests that $UR^{adj.}_{\tau,q,cb}$ should be the preferred test when there is uncertainty over the degree of a low order polynomial trend. The procedure is robust, correctly sized, powerful and easy to implement.

Finally note that the power advantage of $UR^{adj.}_{\tau,q}$ over $UR^{adj.}_{\tau,q,cb}$ is still quite small when no cubic trend component is present. In fact, it seems even smaller than suggested by the asymptotic results of Figure 1.

2.6 Application to Commodity Prices

In this section we demonstrate the practical importance of the theoretical analysis of the previous sections. We apply the tests discussed above to the same (yet updated) set of relative commodity price series analyzed in Harvey et al. (2011). The data consist of indices of primary commodity prices relative to manufactured goods for 24 commodity categories for the period $1900-2007$, measured in logarithms. Thus the time series each consist of 108 observations.

The prices of these commodities have been studied extensively to test downward trends in the relative commodity prices, namely, The Prebisch-Singer hypothesis (Prebisch 1950, Singer 1950). Our interest here, however, is assessing the order of integration of these series, adding another perspective to these studies. We use the data originally compiled by Grilli and Yang (1988) and updated by Pfaffenzeller, Newbold and Rayner (2007). For the QD detrended Dickey-Fuller regressions we use the Perron and Qu (2007) MAIC procedure to choose the lag order, as we did in the simulations in the previous section. For most of the commodities our results match those of Harvey et al. (2011). However, for a few commodities we are able to reject the unit root by allowing a cubic trend. Moreover, by using the union of rejections strategy we do not reverse any of their rejections because of the loss of power when a cubic trend is not present.

Table 2 lists the results of our seven testing procedures for each commodity index. The values of our $DF-QD^{\tau}$ and $DF-QD^{q}$ statistics differ slightly from those of Harvey et al. (2011) because we use 4 additional years of data. However, our 5% rejections do not differ from theirs for these tests nor do they for $UR_{\tau,q}$ and $UR^{adj.}_{\tau,q}$. Now consider the results of the new tests discussed in this paper in columns 3, 6 and 7. For sugar the cubic test, $DF-QD^{cb}$, rejects the unit root despite the fact that the tests used by Harvey et al. (2011) do not. By definition the null is also rejected by the $UR_{\tau,q,cb}$ test, but note that the null is also still rejected by $UR^{adj.}_{\tau,q,cb}$. This shows that our proposed test, $UR^{adj.}_{\tau,q,cb}$, is able to detect stationarity where Harvey et al. (2011) did not detect it.

Of course, a major concern in allowing for a cubic trend in the testing procedure is that the power of it will suffer, i.e., there will be fewer rejections in general across the set of commodities. Indeed, note that for coffee, tea, rice, rubber and timber the cubic test fails to reject although either the linear test, the quadratic test, or both reject the null. Note however, that the adjusted union of rejections strategy is still able to detect stationarity for timber and rice. While the value of the cubic test statistic $DF - QD^{cb}$ is quite low for these two time series the values of the linear and quadratic statistics are large enough to still reject the null while allowing for a cubic trend. Also note that our preferred test, $UR^{adj.}_{\tau,q,cb}$, agrees with $UR^{adj.}_{\tau,q}$ for all five of these goods.

Overall note that the only commodity series for which our proposed statistic, $UR^{adj.}_{\tau,q,cb}$, and that of Harvey et al. (2011), $UR^{adj.}_{\tau,q}$, disagree at the 5% level is sugar. This application shows the practical use of allowing higher degree polynomial trends. The stationarity of the sugar commodity series cannot be detected without allowing a cubic trend. On the other hand this application has shown that the cubic based testing strategy is very reliable in that the power of the test does not suffer too much for those series that do not appear to have a cubic trend; indeed it does not perform any worse than the quadratic union test in this regard.

The commodity price series are plotted in Figure 3 along with the estimated linear, quadratic, and cubic trends. From these we can see that several series clearly exhibit a quadratic trend, as noted in Harvey et al. (2011). We can also see that a cubic trend fits the price series for sugar better than either a linear or quadratic. This distinction seems subtle but as we have seen makes an important difference when testing for a unit root in the detrended series.

2.7 Statistical Arbitrage and Testing for Cointegration

The methodology developed in the previous sections can also be applied to tests for cointegration in multivariate time series. A pair of series is said to be cointegrated if each series individually is $I(1)$ but some linear combination of the two series is $I(0)$. See Engle and Granger (1987) for an early source on the concept of cointegration; see Harris, McCabe and Leybourne (2002) for a more recent discussion of the literature on cointegration. If the two series are presumed to be nonstationary then a test of the null of no cointegration is essentially a unit root test. The residual from the regression of one of the series on the other can be tested for the presence of the unit root. If such a test rejects the unit root null then we consider the null of no cointegration rejected as well. Depending on the statistical properties of the individual series though a simple unit root test may not be appropriate. The cubic trend union of rejections approach developed in Sections 2-5 can be used to account for the possibility that at least one of the two series exhibits trend behavior of a cubic form.

Consider two series, y_{1t} and y_{2t}, each generated according to equations 2.1 and 2.2 with $\rho_1 = \rho_2 = 1$. Furthermore, suppose that $u_{2t} = \alpha_1 u_{1t} + w_t$ where $w_t = \rho_{wT} w_{t-1} + \nu_t$. Then we are interested in the null hypothesis $H_0 : \rho_{wT} = 1$ versus the (local) alternative, $H_1 : \rho_{wT} = 1 - c_w/T, \quad c_w > 0$.

We propose the following procedure for testing this null hypothesis. First, we regress $\mathbf{y}_{2c}$ on $\mathbf{Z}_c$ where $\mathbf{y}_{2c} = (y_{21}, y_{22} - \rho_c y_{21}, \ldots, y_{2T} - \rho_c y_{2,T-1})'$ and $\mathbf{z}_c = (z_1, z_2 - \rho_c z_1, \ldots, z_T - \rho_c z_{T-1})'$ where $z_t = (1, t, t^2, t^3, y_{1t})'$ and $\rho_c = 1 - c/T$. Next, we obtain the residuals from this regression, $\hat{u}_t$. Finally we run the Dickey-Fuller type residual regression

$$\hat{u}_t = \rho \hat{u}_{t-1} + \sum_{j=1}^{p} \phi_j \Delta \hat{u}_{t-j} + e_t$$

and use the t-test for $\rho = 1$. We determine the number of lags, p, according to the MAIC procedure of Ng and Perron (2001) and Perron and Qu (2007). We use $c = 23$ as recommended for the test above.

We suspect that the asymptotic distribution of this test statistic will be of the same form as that given in Lemma 1 since we are not introducing a new regression step but incorporating the regression of y_{2t} on y_{1t} into the first-stage quasi-detrending regression. However, we also ran Monte Carlo simulations to determine the critical values of the cointegration test. The size is largest when there is no time trend at all. So we simulated u_{1t} and u_{2t} according to the model of equation 2.2 with $\rho = 1$ and set $y_{1t} = u_{1t}$ and $y_{2t} = u_{1t} + \gamma u_{2t}$ for various values of $\gamma \neq 0$. The error processes, ε_t were drawn i.i.d. from a standard normal. We used $10,000$ Monte Carlo iterations. For a large sample, $T = 1000$, we found very similar critical values. In particular, the 5% critical value is -3.8.

We apply this test to the 30 stocks of the Dow Jones Industrial Average for a time span of January 3, 2000 - December 8, 2009. We divide the data into five sections each of 500 days, corresponding to about two years worth of business days. Recall that in order for two series to be cointegrated they must each separately be $I(1)$. Moreover, applying the cointegration test above to stationary series is inappropriate and leads to distortions. Hence we employ a pre-test within each period to determine which series are $I(1)$. The pre-test we use is the cubic union test for a unit root described in Section 4. We then apply a union of rejections test to maintain power for series that do not exhibit a cubic trend. This applies the union of rejections strategy described in Section 4 to the cointegration test developed here allowing for a linear, quadratic or cubic trend. We apply the test to all possible pairs of the series. Table 3 reports the results.

We use three different testing strategies as a comparison. The first set of tests only allows for a linear trend, both in the pre-test and the cointegration test. The second set of tests in the table allows for up to a cubic trend, both in the pre-test and the cointegration test. The third set of tests is a hybrid strategy that uses a linear trend test in the pre-testing stage but employs the cubic union of rejections procedure for the cointegration test.

As we discussed both in the asymptotic analysis and the CPI commodity price application, the cubic test can either result in fewer or more rejections than the linear test. The same is true of the cointegration test used here. Suppose a given pair of series is indeed cointegrated. Then if one or both of the series exhibits a cubic trend then the linear test may attribute this trend in the residual to a unit root and conclude that the two series are not cointegrated. On the other hand, if they are cointegrated but both exhibit only a linear time trend then the cubic test may have low power and hence be less likely to detect cointegration.

As can be seen in Table.3, the most dramatic evidence that the cubic test can improve upon the test that assumes only a linear trend is found in the first and last time periods. In the first range the same series are maintained after discarding the stationary series, regardless of which test is used to determine this. Subsequently, the cubic test finds 39 cointegrating pairs, whereas the linear test only finds 22. Indeed, the linear test finds that only two of the series are cointegrated with at least one other, while the cubic test finds that six are. The same pattern holds for the last time period. While one more series is discarded by the cubic pre-test, the cubic cointegration test still detects far more cointegrating pairs: 95 versus 55 for the linear cointegration test. Stated another way, the cubic test finds that 20 of the series are cointegrated with at least one other, while the linear test finds that only 11 are.

The fourth period exhibits the same pattern as well, though to a lesser extent. Notice, however, that the second and third period show the reverse pattern. That is, in both cases the cubic test finds *fewer* cointegrating pairs than the linear cointegration test. By considering the hybrid strategy in the third set of columns in the table it is apparent that this is because the cubic pre-test discards more series and that these series are more likely to be cointegrated with some other series. However, notice too, that for the third time period when performed on the same pairs of time series the cubic and linear cointegration tests detect the same number of cointegration pairs.

In conclusion we find that allowing for a cubic trend when testing for cointegration typically leads to detection of more cointegrating pairs. This is a useful result for the pairs trader who can find more statistical arbitrage opportunities when a cointegrated pair of stocks deviates far enough. An interesting extension to this research would be to design a trading strategy that builds on this model and analyses the economic value of the findings.

2.8 Conclusion

In this paper we studied the behavior of Dickey-Fuller unit root tests of the form studied by Elliott et al. (1996) and Harvey et al. (2011) under a cubic trend. We confirm that the power of these tests suffers dramatically in the presence of an unattended cubic trend, both asymptotically and in finite sample simulations. We also propose a test, $DF - QD^{cb}$, that incorporates a cubic trend into the QD detrending procedure. We show that such a test maintains good power when a cubic trend is in fact present but that the power is low when only a linear or quadratic trend is present. These results mirror those of Harvey et al. (2009b) and Harvey et al. (2011). As these authors suggest, we implement a union of rejections strategy that is able to take advantage of the superior power of the cubic test when a cubic trend is present and the superior power of the quadratic test, $DF - QD^{q}$, (or the linear, $DF - QD^{\tau}$) when the polynomial trend is of a lower degree. Our proposed union of rejections testing strategy compares favorably to that proposed by Harvey et al. (2011).

We also show the practical importance of the proposed testing procedure by applying it to commodity price indices. For one commodity series – sugar

– we are able to reject the null hypothesis that the series exhibits a unit root by allowing for a cubic trend and yet the power does not suffer seriously for those series which do not exhibit a cubic trend. The testing procedure is easy to implement, maintains good power and controls the size. In addition, we show that allowing for a cubic trend, the unit root test can be successfully used in the area of statistical arbitrage, allowing the trader to detect more cointegrating relationships.

Even though the critical-value adjusted union test, $UR^{adj.}_{\tau,q,cb}$, is properly sized and has good overall power in the presence of cubic, quadratic or linear trends, we saw that the unadjusted version, $UR_{\tau,q,cb}$, has superior power asymptotically as the magnitude of the cubic trend gets larger. A potential extension to this research is to develop a modified test strategy through pretesting first for a strong cubic trend. This will allow to use the less conservative unadjusted union test when strong cubic trend is indeed detected, benefiting from the added power of this test over its adjusted counterpart.

Appendix

The proof of Lemma 1 follows. We split up the derivation for each of the three statistics.

We will make use of the following three results derived from similar results in Harvey et al. (2011) and Chapter 17 of Hamilton (1994).

$$T^{-(p+1)} \sum_{t=1} t^p = (1+p)^{-1} + o(1)$$
$$T^{-3/2-p} \sum_{t=2}^{T} t^p u_{t-1} \xrightarrow{d} \sigma \int_0^1 s^p W_c(s) ds$$
$$T^{-1/2-p} \sum_{t=2}^{T} t^p \Delta u_t \xrightarrow{d} \sigma \left\{ W_c(1) - p \int_0^1 s^{p-1} W_c(s) ds \right\}$$

We also will make use of the following result. If $T^{-1/2}\tilde{u}_{\lfloor rT \rfloor} \xrightarrow{d} \sigma J_c(r)$ then the unit root test based on these residuals will converge in distribution to

$$\frac{J_c(1)^2 - 1}{2\sqrt{\int_0^1 J_c(r)^2 dr}}$$

This result follows by standard arguments if $J_c(0) = 0$ and assuming that the estimate of the variance of the error in the residual regression converges in probability to σ^2. For each case below we therefore just derive the asymptotic distribution of $T^{-1/2}\tilde{u}_{\lfloor rT \rfloor}$.

Asymptotic distribution of $DF - QD^{cb}$

The statistic, $DF - QD^{cb}$, is based on the residuals $\tilde{u}_t := y_t - \tilde{\mu} - \tilde{\beta}t - \tilde{\gamma}t^2 - \tilde{\delta}t^3$, where $\tilde{\theta} := (\tilde{\mu}, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta})'$ is obtained by OLS. First, we want to look at the asymptotic distribution of $\tilde{u}_t$ evaluated at $\lfloor rT \rfloor$ instead of t because, as we can see from Harvey et al. (2011), this is what the asymptotic distribution of the test statistic depends on. Next, by the FCLT $T^{-1/2}u_{\lfloor rT \rfloor}$ converges in distribution to $\sigma W_c(r)$, the Ornstein-Uhlenbeck process. This suggests that we need to scale $\tilde{u}_{\lfloor rT \rfloor}$ by $T^{-1/2}$ as well. Hence, plugging in the formula above for y_t, we need to derive the asymptotic distribution of

$$T^{-1/2}\tilde{u}_{\lfloor rT \rfloor} = T^{-1/2}u_{\lfloor rT \rfloor} - T^{-1/2}(\tilde{\mu} - \mu) - T^{1/2}(\tilde{\beta} - \beta)r - T^{3/2}(\tilde{\gamma} - \gamma)r^2 - T^{5/2}(\tilde{\delta} - \delta)r^3$$

since the floor functions can be ignored in the limit.

The main difficulty is to find the asymptotic distribution of the vector of coefficients, $\tilde{\theta}$ which is obtained by OLS. First, notice that

$$\tilde{\theta} := (X'X)^{-1} X'Y$$

where X is a T by 4 matrix with first row $(1, 1, 1, 1)$ and t^{th} row $(1 - \bar{\rho}, t - \bar{\rho}(t-1), t^2 - \bar{\rho}(t-1)^2, t^3 - \bar{\rho}(t-1)^3)$ and Y is a vector with $Y_1 = y_1$ and $Y_t = y_t - \bar{\rho}y_{t-1}$. Then this first means that $Y = X\theta + U$ where $U_1 = u_1$ and $U_t = u_t - \bar{\rho}u_{t-1}$. So $\tilde{\theta} - \theta = (X'X)^{-1}X'U$. Since θ does not appear on the right-hand side of this equation the residuals are invariant to the true value of the parameters and hence we can assume without loss of generality that $\theta = 0$.

Next let $A := X'X$ and $B = X'U$. We want to find the asymptotic distribution of the coefficient estimates scaled by the right powers of T. By mimicking the pattern from the linear and quadratic trends considered in Harvey et al. (2011) we can guess that

$$\tilde{\theta}_T := \begin{pmatrix} T^{-1/2}(\tilde{\mu} - \mu) \\ T^{1/2}(\tilde{\beta} - \beta) \\ T^{3/2}(\tilde{\gamma} - \gamma) \\ T^{5/2}(\tilde{\delta} - \delta) \end{pmatrix} = \begin{pmatrix} a_{11} & T^{-1}a_{12} & T^{-2}a_{13} & T^{-3}a_{14} \\ a_{12} & T^{-1}a_{22} & T^{-2}a_{23} & T^{-3}a_{24} \\ T^{-1}a_{13} & T^{-2}a_{23} & T^{-3}a_{33} & T^{-4}a_{34} \\ T^{-2}a_{14} & T^{-3}a_{24} & T^{-4}a_{34} & T^{-5}a_{44} \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2}b_1 \\ T^{-1/2}b_2 \\ T^{-3/2}b_3 \\ T^{-5/2}b_4 \end{pmatrix}$$

$$:= A_T^{-1} B_T$$

To check this we have to compute the inverse of a 4 by 4 matrix. We do this using the method of minors and find that it is indeed true.

Now the idea is that we can find the asymptotic distribution of $\tilde{\theta}_T$ by studying the behavior of A_T and B_T separately. We only need to deal with a_{j4}, b_3 and b_4. The other entries of A_T and B_T are the same as in Harvey et al. (2011).

Now let X_j denote the j^{th} column of X. Also, recall that $\bar{\rho} = 1 - \bar{c}T^{-1}$.

Then

$$\begin{aligned}
T^{-2}a_{14} &= T^{-2}X_1'X_4 \\
&= T^{-2}\left(1+\sum_{t=2}^{T}(1-\bar{\rho})(t^3-\bar{\rho}(t-1)^3)\right) \\
&= T^{-2}\left(1+\bar{c}T^{-1}\sum_{t=2}^{T}(3t^2-3t+1+\bar{c}T^{-1}(t-1)^3)\right) \\
&= 3\bar{c}T^{-3}\sum_{t=2}^{T}t^2+\bar{c}^2T^{-4}\sum_{t=2}^{T}t^3+o(1) \\
&\to \bar{c}+\bar{c}^2/4
\end{aligned}$$

Next,

$$\begin{aligned}
T^{-3}a_{24} &= T^{-3}X_2'X_4 \\
&= T^{-3}\left(1+\sum_{t=2}^{T}(t-\bar{\rho}(t-1))(t^3-\bar{\rho}(t-1)^3)\right) \\
&= T^{-3}\left(1+\sum_{t=2}^{T}(1+\bar{c}T^{-1}(t-1))(3t^2-3t+1+\bar{c}T^{-1}(t-1)^3)\right) \\
&= 3T^{-3}\sum_{t=2}^{T}t^2+3\bar{c}T^{-4}\sum_{t=2}^{T}t^3+\bar{c}T^{-4}\sum_{t=2}^{T}t^3+\bar{c}^2T^{-5}\sum_{t=2}^{T}t^4+o(1) \\
&\to 1+\bar{c}+\bar{c}^2/5=\pi_4
\end{aligned}$$

Next,

$$
\begin{aligned}
T^{-4}a_{34} &= T^{-4}X_3'X_4 \\
&= T^{-4}\left(1+\sum_{t=2}^{T}(t^2-\bar{\rho}(t-1)^2)(t^3-\bar{\rho}(t-1)^3)\right) \\
&= T^{-4}\left(1+\sum_{t=2}^{T}(2t-1+\bar{c}T^{-1}(t-1)^2)(3t^2-3t+1+\bar{c}T^{-1}(t-1)^3)\right) \\
&= 6T^{-4}\sum_{t=2}^{T}t^3+3\bar{c}T^{-5}\sum_{t=2}^{T}t^4+2\bar{c}T^{-5}\sum_{t=2}^{T}t^4+\bar{c}^2T^{-6}\sum_{t=2}^{T}t^5+o(1) \\
&\to 3/2+\bar{c}+\bar{c}^2/6=\pi_5
\end{aligned}
$$

Finally,

$$
\begin{aligned}
T^{-5}a_{44} &= T^{-5}X_4'X_4 \\
&= T^{-5}\left(1+\sum_{t=2}^{T}(t^3-\bar{\rho}(t-1)^3)^2\right) \\
&= T^{-5}\left(1+\sum_{t=2}^{T}(3t^2-3t+1+\bar{c}T^{-1}(t-1)^3)^2\right) \\
&= 9T^{-5}\sum_{t=2}^{T}t^4+6\bar{c}T^{-6}\sum_{t=2}^{T}t^5+\bar{c}^2T^{-7}\sum_{t=2}^{T}t^6+o(1) \\
&\to 9/5+\bar{c}+\bar{c}^2/7 := \pi_6
\end{aligned}
$$

Lastly consider b_3 and b_4. Using them we find,

$$T^{-3/2}b_3 = 2T^{-3/2}\sum_{t=2}^{T} t\Delta u_t + 2\bar{c}T^{-5/2}\sum_{t=2}^{T} tu_{t-1} + \bar{c}T^{-5/2}\sum_{t=2}^{T} t^2\Delta u_t + \bar{c}^2T^{-7/2}\sum_{t=2}^{T} t^2 u_{t-1}$$
$$\xrightarrow{d} 2\sigma\left\{W_c(1) - \int_0^1 W_c(s)ds\right\} + 2\bar{c}\sigma\int_0^1 sW_c(s)ds + \bar{c}\sigma\left\{W_c(s) - 2\int_0^1 sW_c(s)ds\right\} +$$
$$\bar{c}^2\sigma\int_0^1 s^2W_c(s)ds$$
$$= \sigma\left\{(2+\bar{c})W_c(1) - 2\int_0^1 W_c(s)ds + \bar{c}^2\int_0^1 s^2W_c(s)ds\right\} := \sigma M_2$$

Note that this is the same asymptotic distribution for $T^{-3/2}b_3$ that is derived in Harvey et al. (2011). Next,

$$T^{-5/2}b_4 = 3T^{-5/2}\sum_{t=2}^{T} t^2\Delta u_t + 3\bar{c}T^{-7/2}\sum_{t=2}^{T} t^2u_{t-1} + \bar{c}T^{-7/2}\sum_{t=2}^{T} t^3\Delta u_t +$$
$$\bar{c}^2T^{-9/2}\sum_{t=2}^{T} t^3u_{t-1} + o_p(T^{5/2})$$
$$\xrightarrow{d} 3\sigma\left\{W_c(1) - 2\int_0^1 sW_c(s)ds\right\} + 3\bar{c}\sigma\int_0^1 s^2W_c(s)ds +$$
$$\bar{c}\sigma\left\{W_c(s) - 3\int_0^1 s^2W_c(s)ds\right\} + \bar{c}^2\sigma\int_0^1 s^3W_c(s)ds$$
$$= \sigma\{(3+\bar{c})W_c(1) - 6\int_0^1 sW_c(s)ds + \bar{c}^2\int_0^1 s^3W_c(s)ds\} := \sigma M_3$$

So, combining these results,

$$\begin{pmatrix} T^{-1/2}\tilde{\mu} \\ T^{1/2}\tilde{\beta} \\ T^{3/2}\tilde{\gamma} \\ T^{5/2}\tilde{\delta} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1+\bar{c}+\bar{c}^2/2 & \pi_1 & \pi_2 & \pi_4 \\ \bar{c}+\bar{c}^2/3 & \pi_2 & \pi_3 & \pi_5 \\ \bar{c}+\bar{c}^2/4 & \pi_4 & \pi_5 & \pi_6 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \sigma M_1 \\ \sigma M_2 \\ \sigma M_3 \end{pmatrix}$$

$$= \sigma \begin{pmatrix} 0 \\ \Pi^{-1} \begin{pmatrix} M_1 \\ M_2 \\ M_3 \end{pmatrix} \end{pmatrix}$$

where

$$\Pi = \begin{pmatrix} \pi_1 & \pi_2 & \pi_4 \\ \pi_2 & \pi_3 & \pi_5 \\ \pi_4 & \pi_5 & \pi_6 \end{pmatrix}$$

Plugging this back into the formula for $T^{-1/2}\tilde{u}_{\lfloor rT \rfloor}$ we obtain $J_c^{cb,\bar{c}_{cb}}(r)$ defined in the theorem.

Asymptotic distribution of $DF - QD^q$

We set $\mu = \beta = \gamma = 0$ since $DF - QD^q$ is invariant to these parameters. Hence $y_t = \kappa_2 T^{-5/2} t^3 + u_t$. The statistic is based on the residuals $\tilde{u}_t := y_t - \tilde{\mu} - \tilde{\beta} t - \tilde{\gamma} t^2$.

Hence we need to derive the asymptotic distribution of

$$T^{-1/2}\tilde{u}_{\lfloor rT \rfloor} = T^{-1/2}u_{\lfloor rT \rfloor} + \kappa_2 r^3 - T^{-1/2}\tilde{\mu} - T^{1/2}\tilde{\beta}r - T^{3/2}\tilde{\gamma}r^2$$

The main difficulty is to find the asymptotic distribution of the vector of coefficients, $(\tilde{\mu}, \tilde{\beta}, \tilde{\gamma})'$ which is obtained by OLS. Notice that

$$\begin{pmatrix} T^{-1/2}\tilde{\mu} \\ T^{1/2}\tilde{\beta} \\ T^{3/2}\tilde{\gamma} \end{pmatrix} = \begin{pmatrix} a_{11} & T^{-1}a_{12} & T^{-2}a_{13} \\ a_{12} & T^{-1}a_{22} & T^{-2}a_{23} \\ T^{-1}a_{13} & T^{-2}a_{23} & T^{-3}a_{33} \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2}b_1 \\ T^{-1/2}b_2 \\ T^{-3/2}b_3 \end{pmatrix}$$

The terms a_{ij} are the same as in HLT. So we only need to deal with b_1, b_2, b_3.

$$b_1 = y_1 + \bar{c}T^{-1}(y_T - y_1) + \bar{c}^2 T^{-2}\sum_{t=2}^{T} y_{t-1}$$
$$b_2 = y_T + \bar{c}T^{-1}\sum_{t=2}^{T} y_t + \bar{c}T^{-1}\sum_{t=2}^{T} t\Delta y_t + \bar{c}^2 T^{-2}\sum_{t=2}^{T} t y_{t-1} + o_p(T^{1/2})$$
$$b_3 = 2\sum_{t=2}^{T} t\Delta y_t + 2\bar{c}T^{-1}\sum_{t=2}^{T} t y_{t-1} + \bar{c}T^{-1}\sum_{t=2}^{T} t^2\Delta y_t + \bar{c}^2 T^{-2}\sum_{t=2}^{T} t^2 y_{t-1} + o_p(T^{3/2})$$

We will make use of the fact that $T^{-(p+1)}\sum_{t=1} t^p = (1+p)^{-1} + o(1)$. For example, $\sum_{t=1}^{T} t = \frac{1}{2}T(T+1)$, so dividing by T^2 we get $\frac{1}{2}(1+T^{-1}) \to \frac{1}{2}$.

$$T^{-1/2}b_1 = T^{-1/2}u_1 + \kappa_2 T^{-3} + \bar{c}T^{-3/2}(u_T - u_1 + \kappa_2 T^{1/2} - \kappa_2 T^{-5/2}) + \bar{c}^2 T^{-5}\kappa_2 \sum_{t=2}^{T} t^3 +$$
$$\bar{c}^2 T^{-5/2}\sum_{t=2}^{T} u_{t-1} \to_p 0$$

Next consider b_2.

$$T^{-1/2}b_2 = \kappa_2 + T^{-1/2}u_T + \bar{c}T^{-4}\kappa_2\sum_{t=2}^{T}t^3 + \bar{c}T^{-3/2}\sum_{t=2}^{T}u_{t-1} + \bar{c}T^{-4}\kappa_2\sum_{t=2}^{T}t(t^3-(t-1)^3)$$
$$+ \bar{c}T^{-3/2}\sum_{t=2}^{T}t\Delta u_t + \bar{c}^2T^{-5}\kappa_2\sum_{t=2}^{T}t^4 + \bar{c}^2T^{-5/2}\sum_{t=2}^{T}tu_{t-1} + o_p(T^{1/2})$$
$$\xrightarrow{d} \sigma\left\{(1+\bar{c})W_c(1) + \bar{c}^2\int_0^1 sW_c(s)ds + \kappa_2^*\left(1+\bar{c}+\frac{\bar{c}^2}{5}\right)\right\}$$

where $\kappa_2^* := \sigma^{-1}\kappa_2$.

Lastly consider b_3.

$$T^{-3/2}b_3 = 2T^{-3/2}\sum_{t=2}^{T}t\Delta u_t + 2\bar{c}T^{-5/2}\sum_{t=2}^{T}tu_{t-1} + \bar{c}T^{-5/2}\sum_{t=2}^{T}t^2\Delta u_t + \bar{c}^2T^{-7/2}\sum_{t=2}^{T}t^2u_{t-1}$$
$$+ 2T^{-4}\kappa_2\sum_{t=2}^{T}t(t^3-(t-1)^3) + 2\bar{c}T^{-5}\kappa_2\sum_{t=2}^{T}t^4 + \bar{c}T^{-5}\kappa_2\sum_{t=2}^{T}t^2(t^3-(t-1)^3)+$$
$$\bar{c}^2T^{-6}\kappa_2\sum_{t=2}^{T}t^5$$
$$\xrightarrow{d} \sigma\left\{(2+\bar{c})W_c(1) - 2\int_0^1 W_c(s)ds + \bar{c}^2\int_0^1 s^2W_c(s)ds + \kappa_2^*\left(\frac{3}{2}+\bar{c}+\frac{\bar{c}^2}{6}\right)\right\}$$

Combining these results with the asymptotic forms of a_{ij} derived in Harvey et al. (2011) we get

$$\begin{pmatrix} T^{-1/2}\tilde{\mu} \\ T^{1/2}\tilde{\beta} \\ T^{3/2}\tilde{\gamma} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} 0 \\ \sigma d^{-1}(\pi_3M_1 - \pi_2M_2 + \kappa_2^*(\pi_3\pi_4 - \pi_2\pi_5)) \\ \sigma d^{-1}(\pi_1M_2 - \pi_2M_1 + \kappa_2^*(\pi_1\pi_5 - \pi_2\pi_4)) \end{pmatrix}$$

where everything is defined in Harvey et al. (2011) except

$$\pi_4 = 1 + \bar{c} + \frac{\bar{c}^2}{5}$$
$$\pi_5 = \frac{3}{2} + \bar{c} + \frac{\bar{c}^2}{6}$$

Lastly, this implies that

$$T^{-1/2}\tilde{u}_{\lfloor rT \rfloor} \xrightarrow{d} \sigma\left\{W_c(r) - rd^{-1}(\pi_3 M_1 - \pi_2 M_2) - r^2 d^{-1}(\pi_1 M_2 - \pi_2 M_1)\right.$$
$$\left. + \kappa_2^*(r^3 - rd^{-1}(\pi_3\pi_4 - \pi_2\pi_5) - r^2 d^{-1}(\pi_1\pi_5 - \pi_2\pi_4)\right\}$$

Asymptotic distribution of $DF - QD^\tau$

We set $\mu = \beta = 0$ since $DF - QD^\tau$ is invariant to these parameters. Hence $y_t = \kappa_1 T^{-3/2} t^2 + \kappa_2 T^{-5/2} t^3 + u_t$. The statistic is based on the residuals $\tilde{u}_t := y_t - \tilde{\mu} - \tilde{\beta} t$. Hence we need to derive the asymptotic distribution of

$$T^{-1/2}\tilde{u}_{\lfloor rT \rfloor} = T^{-1/2} u_{\lfloor rT \rfloor} + \kappa_1 r^2 + \kappa_2 r^3 - T^{-1/2}\tilde{\mu} - T^{1/2}\tilde{\beta} r$$

So we have to find the asymptotic distribution of the vector of coefficients, $(\tilde{\mu}, \tilde{\beta})'$ which is obtained by OLS. First notice that

$$\begin{pmatrix} T^{-1/2}\tilde{\mu} \\ T^{1/2}\tilde{\beta} \end{pmatrix} = \begin{pmatrix} a_{11} & T^{-1}a_{12} \\ a_{12} & T^{-1}a_{22} \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2}b_1 \\ T^{-1/2}b_2 \end{pmatrix}$$

The terms a_{ij} are the same as in HLT. As in HLT, $T^{-1/2}b_1 \to 0$. Next,

$$\begin{aligned} b_2 &= y_1 + \sum_{t=2}^{T}(y_t - \bar{\rho}y_{t-1})(t - \bar{\rho}(t-1)) \\ &= y_1 + \sum_{t=2}^{T}(\Delta y_t + \bar{c}T^{-1}y_{t-1})(1 + \bar{c}T^{-1}(t-1)) \end{aligned}$$

So

$$\begin{aligned} T^{-1/2}b_2 &= o_p(1) + T^{-1/2}y_T + \bar{c}T^{-3/2}\sum_{t=1}^{T} y_t + \bar{c}T^{-3/2}\sum_{t=1}^{T} t\Delta y_t + \bar{c}^2T^{-5/2}\sum_{t=1}^{T} ty_t \\ &= o_p(1) + \kappa_1 + \kappa_2 + T^{-1/2}u_T + 3\bar{c}\kappa_1 T^{-3}\sum_{t=1}^{T} t^2 + 4\bar{c}\kappa_2 T^{-4}\sum_{t=1}^{T} t^3 + \bar{c}T^{-3/2}\sum_{t=1}^{T} u_t \\ &+ \bar{c}T^{-3/2}\sum_{t=1}^{T} t\Delta u_t + \bar{c}^2\kappa_1 T^{-4}\sum_{t=1}^{T} t^3 + \bar{c}^2\kappa_2 T^{-5}\sum_{t=1}^{T} t^4 + \bar{c}^2 T^{-5/2}\sum_{t=1}^{T} tu_t \\ &\xrightarrow{d} \sigma M_{1,\bar{c}} + \kappa_1(1 + \bar{c} + \bar{c}^2/4) + \kappa_2(1 + \bar{c} + \bar{c}^2/5) \end{aligned}$$

where $M_{1,\bar{c}}$ is defined in Harvey et al. (2011) as $(1+\bar{c})W_c(1) + \bar{c}^2\int_0^1 sW_c(s)ds$. Combining these results

$$\begin{aligned} \begin{pmatrix} T^{-1/2}\tilde{\mu} \\ T^{1/2}\tilde{\beta} \end{pmatrix} &\xrightarrow{d} \begin{pmatrix} 1 & 0 \\ 1 + \bar{c} + \bar{c}^2/2 & \pi_{1,\bar{c}} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \sigma M_{1,\bar{c}} + \kappa_1(1 + \bar{c} + \bar{c}^2/4) + \kappa_2(1 + \bar{c} + \bar{c}^2/5) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \pi_{1,\bar{c}}^{-1}(\sigma M_{1,\bar{c}} + \kappa_1(1 + \bar{c} + \bar{c}^2/4) + \kappa_2(1 + \bar{c} + \bar{c}^2/5)) \end{pmatrix} \end{aligned}$$

Finally we can plug this into the expression above to get the asymptotic distribution of $T^{-1/2}\tilde{u}_{\lfloor rT \rfloor}$.

Table 2. Unit Root Tests on Commodity Price Data

Commodity	DF-DQτ	DF-DQq	Df-DQcb	UR τ-q	UR τ-q Adj	UR τ-q-cb	UR τ-q-cb Adj
Coffee	**-3.000**	**-3.628**	-3.859	1	0	1	0
Cocoa	-2.110	-2.163	-2.924	0	0	0	0
Tea	-2.215	**-3.568**	-3.743	1	0	1	0
Rice	-1.790	**-3.809**	-3.854	1	1	1	1
Wheat	**-3.340**	**-3.828**	**-4.040**	1	1	1	1
Maize	-0.704	**-4.946**	**-5.258**	1	1	1	1
Sugar	**-2.903**	-2.906	**-4.910**	1	0	1	1
Beef	**-2.975**	-2.991	-3.339	1	0	1	0
Lamb	**-3.085**	-3.091	-3.098	1	1	1	1
Banana	-1.598	-2.376	-3.216	0	0	0	0
Palmoil	-2.792	**-4.291**	**-4.320**	1	1	1	1
Cotton	-1.239	-2.459	-2.451	0	0	0	0
Jute	-0.913	-1.801	-1.835	0	0	0	0
Wool	-1.405	-2.162	-2.892	0	0	0	0
Hides	-1.616	-3.271	-3.676	0	0	0	0
Tobacco	-0.738	**-4.124**	**-4.231**	1	1	1	1
Rubber	**-2.858**	**-3.483**	-3.489	1	0	1	0
Timber	**-3.519**	**-3.780**	-3.795	1	1	1	1
Copper	-2.222	-2.715	-2.722	0	0	0	0
Aluminum	-2.528	-3.345	-3.365	0	0	0	0
Tin	-2.569	-2.738	-2.832	0	0	0	0
Silver	-1.859	-2.237	-2.761	0	0	0	0
Lead	-2.419	-2.437	-2.553	0	0	0	0
Zinc	**-3.980**	**-4.157**	**-4.181**	1	1	1	1

Table 3. Cointegration tests with trend uncertainty

period	linear				cubic				linear/cubic			
	cointegrating pairs detected	# series	# pairs tested	% pairs detected	cointegrating pairs detected	# series	# pairs tested	% pairs detected	cointegrating pairs detected	# series	# pairs tested	% pairs detected
1	22	26	650	0.034	39	26	650	0.060	45	26	650	0.069
2	36	27	702	0.051	33	25	600	0.055	65	27	702	0.093
3	31	28	756	0.041	21	29	812	0.026	31	28	756	0.041
4	17	25	600	0.028	24	26	650	0.037	24	25	600	0.040
5	55	30	870	0.063	95	29	812	0.117	103	30	870	0.118

Figure 1. Asymptotic size and local power of the linear, quadratic and cubic under cubic

Model: $y_t = \mu + \beta t + \kappa_1 T^{-3/2} t^2 + \kappa_2 T^{-5/2} t^3 + u_t$

(Monte Carlo=500, T=1000)

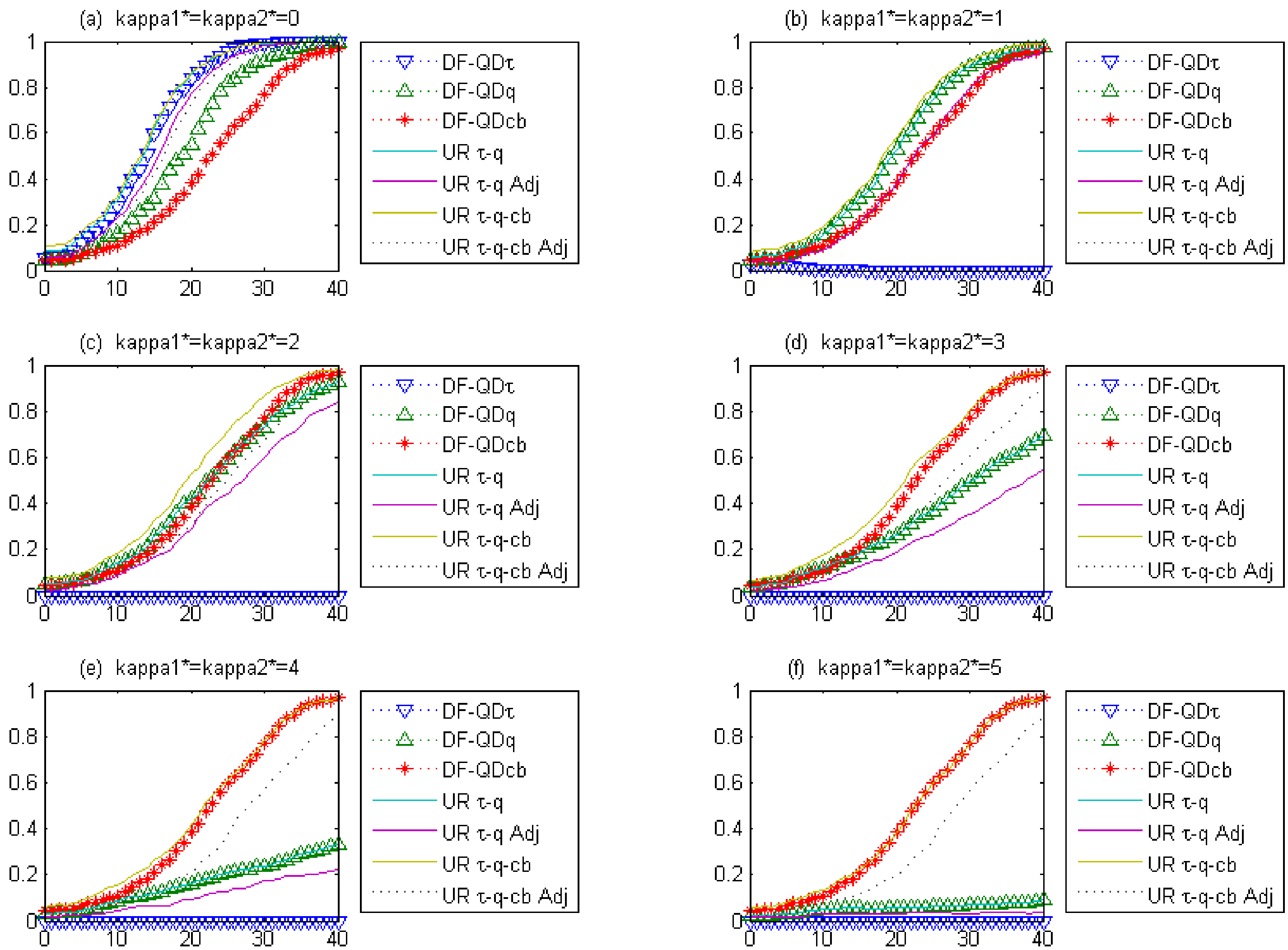


Figure 2. Finite samples size and local power of the linear, quadratic and cubic under cubic

Model: $y_t = \mu + \beta t + \kappa_1 T^{-3/2} t^2 + \kappa_2 T^{-5/2} t^3 + u_t$

(Monte Carlo=1000, T=150)

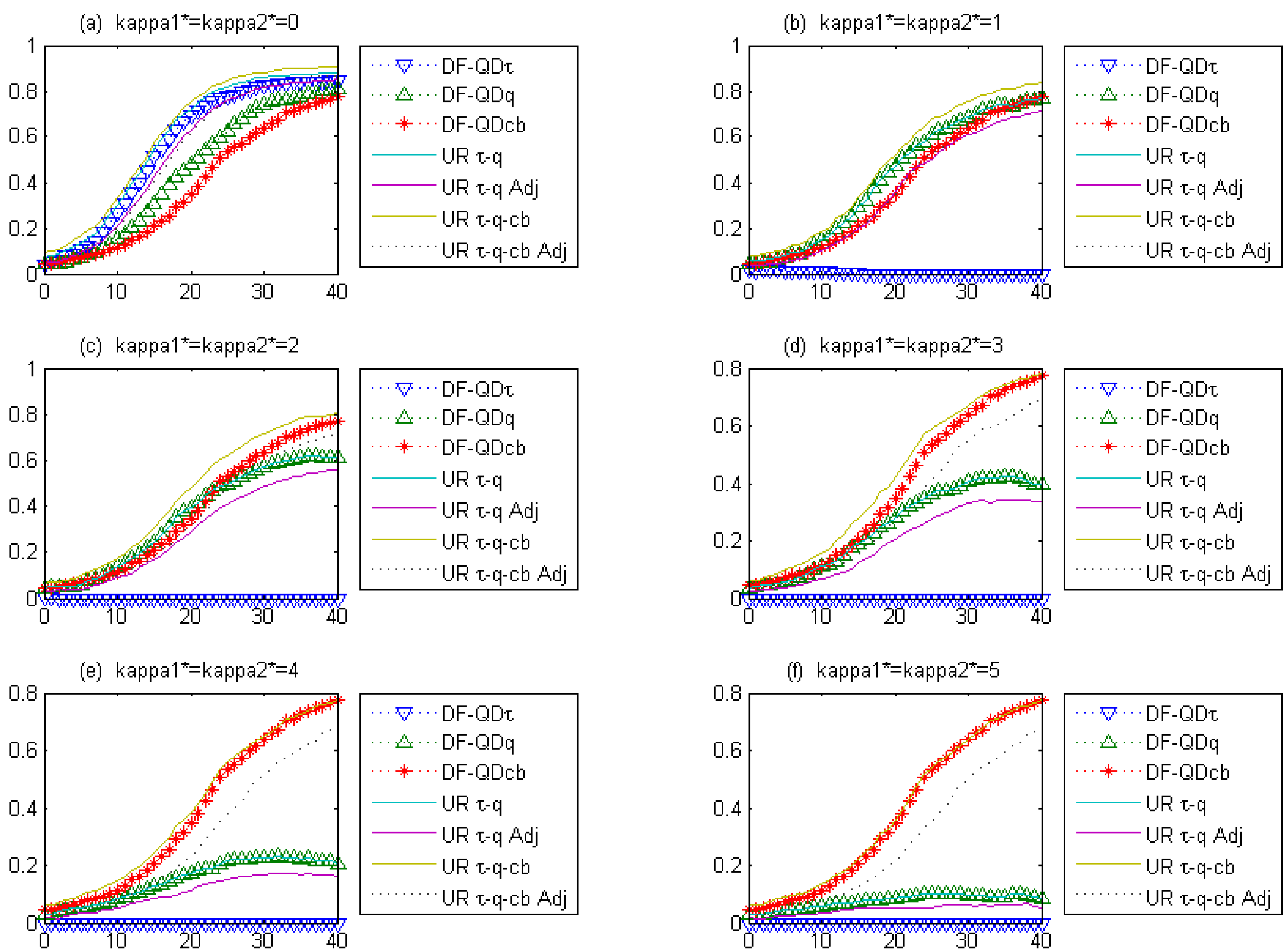


Figure 3(a). Relative primary commodity price series and fitted deterministic components

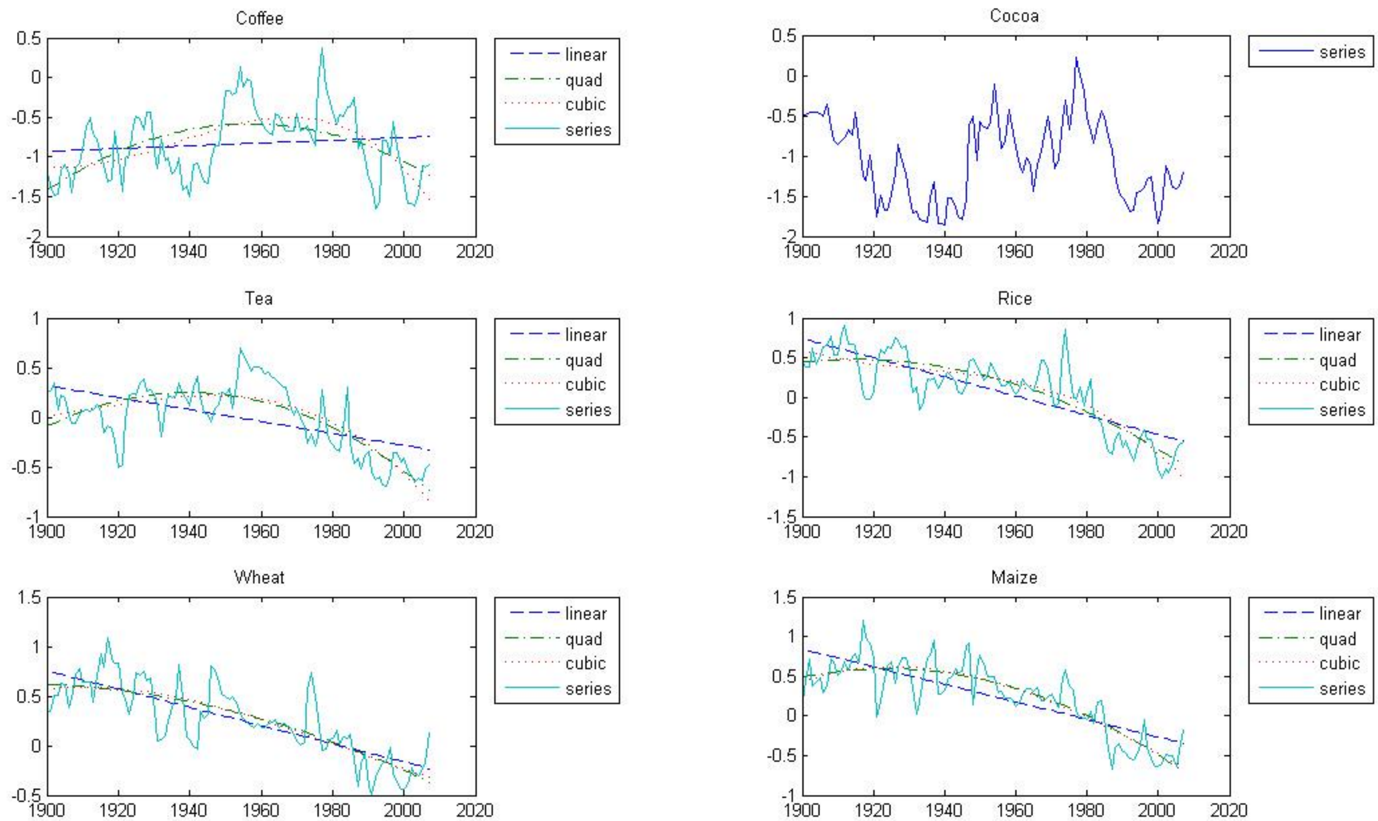


Figure 3(b). Relative primary commodity price series and fitted deterministic components

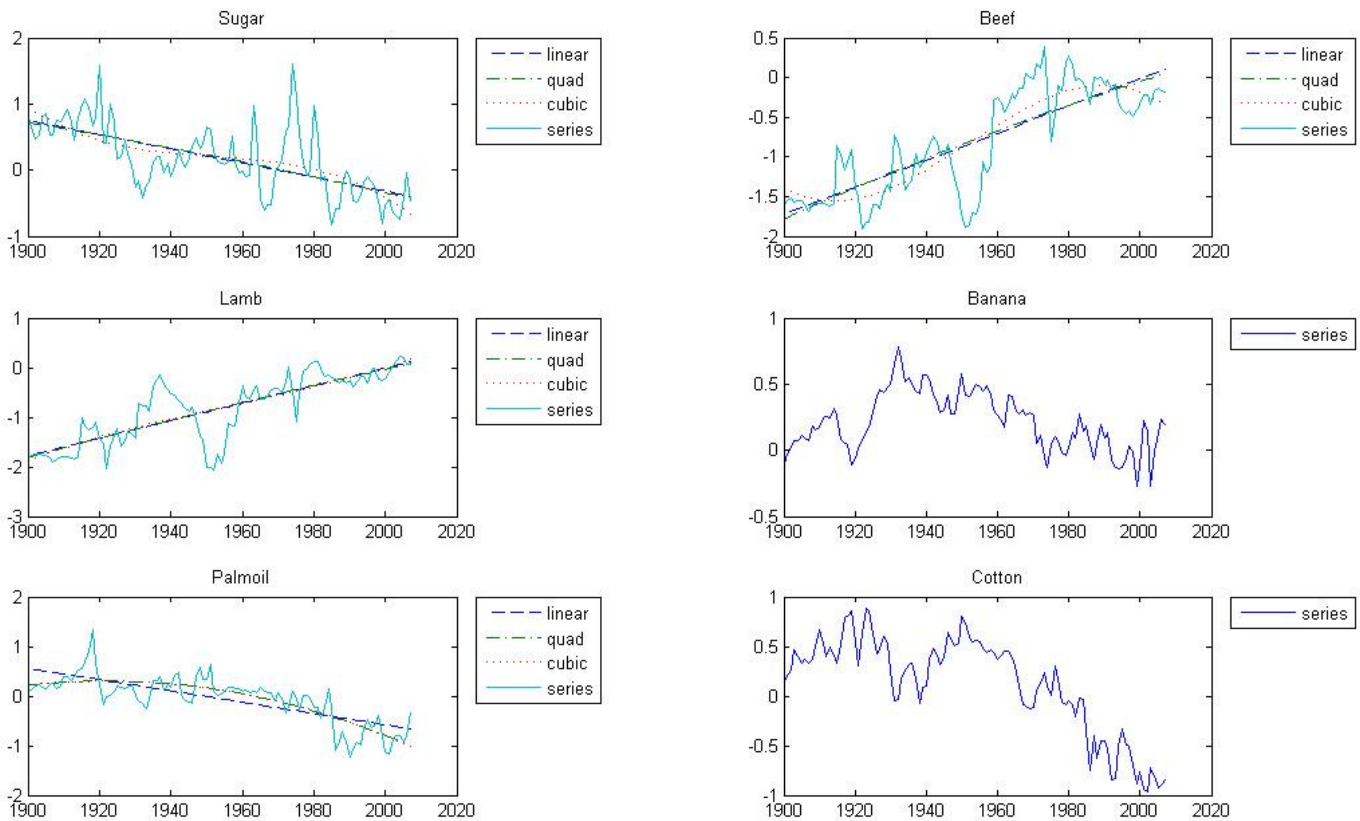


Figure 3(c). Relative primary commodity price series and fitted deterministic components

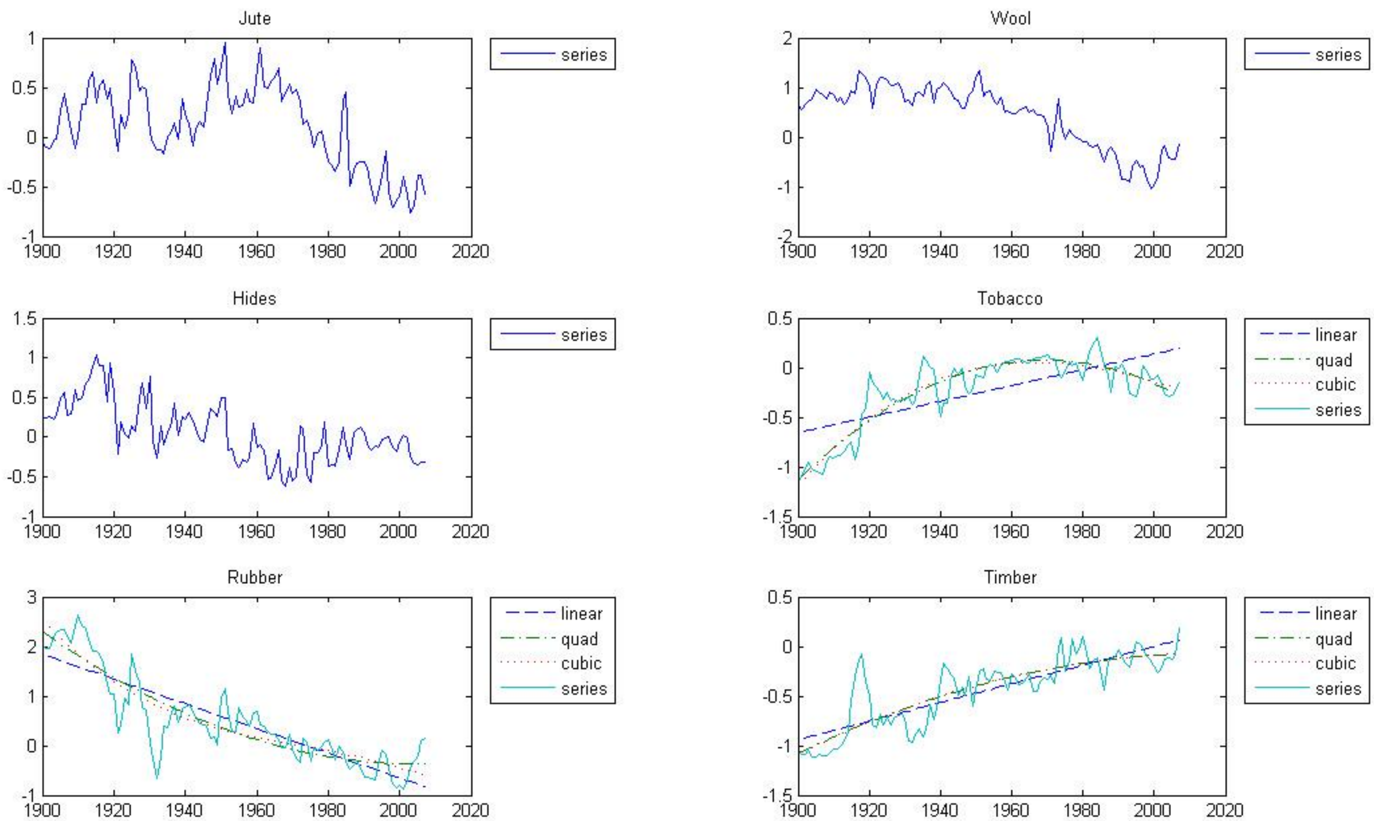
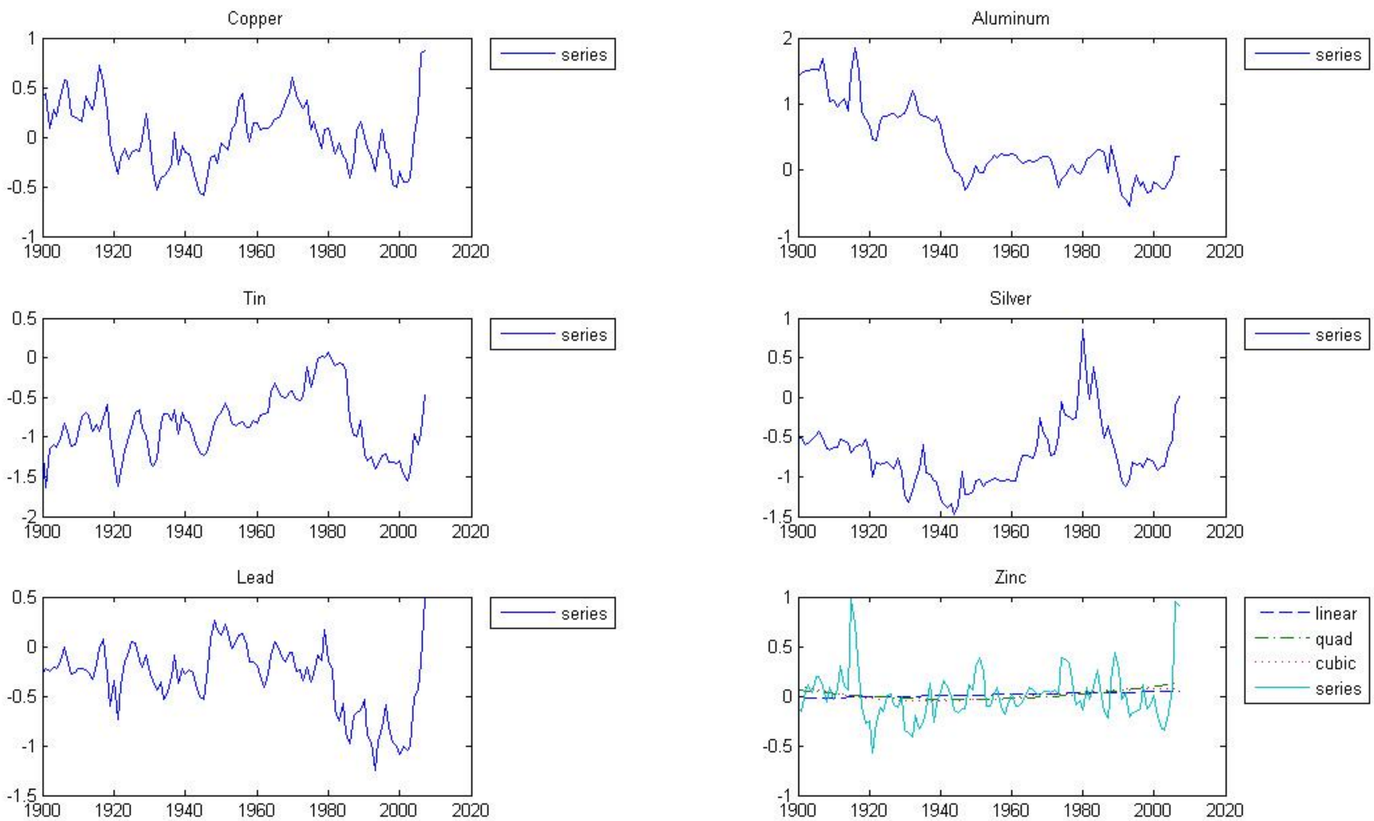


Figure 3(d). Relative primary commodity price series and fitted deterministic components



Chapter 3

The robustness of a panel stationarity test to nonstationary volatility

Harris, Leybourne and McCabe (2005) propose a test for stationarity in panel time series data that allows for arbitrary cross-sectional dependence while also treating the time series dynamics nonparametrically. They do not, however, account for the possibility of time series heteroskedasticity. We explore the behavior of their test in finite samples for ARMA models when the underlying innovation has a time-varying variance through a series of Monte Carlo exercises and discover the unexpected result that the test is fairly robust without correcting for the heteroskedasticity. We find that the uncorrected statistic of Harris et al. (2005) that does not account for small sample bias arising from estimation of the regression coefficients is still conservative in our simulations which allow for non-constant volatility. That is, we find that the size is always less than the nominal size. Second, we find that the bias-corrected version of their test is typically over-sized but the distortion is not substantial in any of our simulations. Third, we find that in general the unattended nonstationary volatility does not adversely affect the power of either test.

3.1 Introduction

The proliferation of long time series on many macroeconomic variables recently has led to the development of new ways to test for unit roots, or conversely test for stationarity, in panels. As long as the separate series are not highly correlated the use of panels should lead to test with higher power (see, inter alia, O'Connell (1998); Maddala and Wu (1999); Hadri (2000); Choi (2001); Choi (2006); Im, Pesaran and Shin (2003); and Harris et al. (2005)).

One key problem with some of the proposed tests is that they assume cross-sectional independence. Harris et al. (2005) suggest a way to allow arbitrary cross-sectional dependence while still allowing for a wide range of time series dynamics. Their test is essentially the sum of the lag-k estimate autocorrelations across the N series. By allowing k to increase with time in the asymptotic analysis they are able to show that asymptotically this test can handle all sorts of time series behavior. This sum of autocorrelations is then asymptotically normal under standard conditions and the only other component needed to perform a valid test for stationarity is an estimate of the asymptotic variance. Their test is then robust to cross-sectional dependence because they use an estimate of the asymptotic variance that allows for arbitrary dependence between the series. This estimate relies on a consistent estimate of the long-run variance of the series $\sum_{i=1}^{N} e_{it}e_{i,t-k}$. This series essentially describes the cross-sectional lag-k autocorrelation at each period t.

While allowing for arbitrary cross-sectional dependence and serial correlation, Harris et al. (2005) assumes that the underlying process in each series is stationary. However, the importance of accounting for non-stationary unconditional volatility has recently been recognized in a growing literature. In particular, a large body of recent work has shown that the unconditional volatility of the processes driving many macroeconomic time series declined over the last quarter of the last century (see, e.g., the literature review in Cavaliere and Taylor, 2008). Sensier and Dijk (2004) find wide spread evidence of non-constant volatility in the Stock and Watson (1999) dataset. Moreover, unattended time-varying volatility has been found to produce significant size distortions in standard unit root and stationary tests (Kim, Leybourne and Newbold 2002, Busetti and Taylor 2003, Cavaliere 2004, Cavaliere and Taylor 2005, Cavaliere and Taylor 2007, Cavaliere and Taylor 2008), as well as in tests of level breaks (Daihes 2011).

Because of these concerns we conduct a series of Monte Carlo experiments to assess the distortion caused by non-constant volatility in the time series. One would expect that a heteroskedasticity-robust variance estimator is necessary for the test to be valid. Surprisingly we find little distortion in the size of the test without such a correction. In Section 2 we lay out the model considered by Harris et al. (2005) and define their proposed estimator. In Section 3 we describe our simulations and report the results. Section 4 concludes.

3.2 The model

Suppose we observe N time series of length T, $\{y_{it}\}$. The model, or data generating process, we assume is

$$y_{it} = \beta_i' x_{it} + e_{it} \tag{3.1}$$

$$e_{it} = \phi_i e_{i,t-1} + \varepsilon_{it}, t = 1, \ldots, T \tag{3.2}$$

We wish to test the null hypothesis

$$H_0 : \quad \phi_i < 1 \text{ for all } i$$

against the alternative that $\phi_i = 1$ for some i.

Harris et al. (2005) assume that $\boldsymbol{\varepsilon_t} = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})$ satisfies the following assumption.

Assumption 3.1. *$\boldsymbol{\varepsilon_t}$ is a $N \times 1$ vector of fixed dimension generated by*

$$\boldsymbol{\varepsilon_t} = \mathbf{A}(L)\boldsymbol{\xi_t},$$

where $\mathbf{A}(L) = \sum_{j=0}^{\infty} \mathbf{A}_j L^j$ and $\mathbf{A}_j$ and $\boldsymbol{\xi_t}$ satisfy

(i). $A_0 = I_N$

(ii). $\sum_{j=0}^{\infty} j^2 tr(\mathbf{A}_j' \mathbf{A}_j) < \infty$

(iii). $\mathbf{A}(1)$ *has full rank*

(iv). $\{\boldsymbol{\xi_t}, \mathcal{F}_t\}$ *is a martingale difference sequence where* $\mathcal{F}_t = \sigma\{\boldsymbol{\xi_{t-j}}, j \geq 0\}$

(v). $E(\boldsymbol{\xi_t}\boldsymbol{\xi_t}' \mid \mathcal{F}_{t-1}) = \Sigma$ *almost surely, for all* t

(vi). $||E(\boldsymbol{\xi_t}\boldsymbol{\xi_t}' \otimes \boldsymbol{\xi_t}\boldsymbol{\xi_t}' \mid \mathcal{F}_{t-1})|| < \kappa < \infty$ *almost surely, for all* t[1]

This assumption allows for substantial dependence across i and serial correlation across t as well as heteroskedasticity across i. The one thing it does not allow is heteroskedasticity across t. Assumption 1*(v)* restricts the driving process $\boldsymbol{\xi_t}$ to be stationary. Alternatively we might assume that $\xi_t = \Sigma_t \nu_t$ for a non-stochastic, time-varying volatility series, Σ_t, and a process ν_t that satisfies Assumptions 1*(iv)*-1*(vi)*. We find in a series of Monte Carlo exercises that their procedure is robust to this more general specification. Specifically, we simulate processes with a jump in volatility, i.e., $\Sigma_t = \sigma_t I_N$ for a scalar $\sigma_t = \sigma_0 + \mathbb{I}(t > \lfloor .5T \rfloor)(\sigma_1 - \sigma_0)$. Harris et al. (2005) also assume the following about the regressors x_{it}.

Assumption 3.2. *For each* i *there exists* D_{iT} *such that (i)* $D_{iT}^{-1} x_{i\lfloor \tau T \rfloor} \rightarrow X_i(\tau) < \infty$*, uniformly in* τ *and (ii)* $T^{-1} \sum_{t=1}^{T} D_{iT}^{-1} x_{it} \left(D_{iT}^{-1} x_{it}\right)' \rightarrow \int_0^1 X_i(\tau) X_i(\tau)' d\tau > 0$.

This assumptions requires the deterministic regressors to satisfy a very mild set of restrictions regarding their limiting behavior. Note that for example, any polynomial trend will satisfy this assumption. If $x_{it} = t^p$ for some positive integer p then it is satisfied with $D_{iT} = T^p$ and $X_i(\tau) = \tau^p$.

[1] Where $||W||$ denotes $\sqrt{tr(W'W)}$.

3.2.1 The estimator

To motivate the test statistic suppose $y_t = \rho y_{t-1} + e_t$ where e_t is white noise with variance σ^2. Then if $\rho < 1$,

$$E\left(T^{-1/2}\sum_{t=k+1}^{T} y_t y_{t-k}\right) \approx T^{1/2}\sigma^2\rho^k/(1-\rho^2)$$

On the other hand, when $\rho = 1$, the same expectation is approximately $T^{3/2}\sigma^2$ (Harris et al. 2005). Note that using this statistic for $k = 1$ requires estimation of ρ as well as σ^2. More generally for fixed k such a statistic requires specification and estimation of the time series behavior of the data. However, as documented by (Harris et al. 2005), if k is indexed by T so that it is increasing in T but $o(T)$ then the right-hand side of the above expression converges to 0 while $T^{3/2}\sigma^2 \to \infty$[2].

We start by residualizing and normalizing each series. Let $\hat{e}_{it}$ denote the OLS residual and let $\tilde{e}_{it}$ denote $\hat{e}_{it}/(T^{-1}\sum_{t=1}^{T}\hat{e}_{it}^2)$. Next define

$$S_k = T^{-1/2}\sum_{i=1}^{N}\sum_{t=k+1}^{T}\tilde{e}_{i,t}\tilde{e}_{i,t-k}$$

This is the statistic suggested above for a univariate time series aggregated over the N different series. Next define

$$\hat{\Gamma}_j(a_t) = T^{-1}\sum_{t=j+1}^{T} a_t a'_{t-j}$$

[2]To see this, suppose $k = T^\delta$ for $0 < \delta < 1$. Then $T^{1/2}\sigma^2\rho^k/(1-\rho^2) = \sigma^2/(1-\rho^2)\exp((1/2)\ln(T) + T^\delta\ln(\rho))$. Since $\rho < 1$ and because $\ln(T)$ converges slower than any power of T it follows that $(1/2)\ln(T) + T^\delta\ln(\rho) \to -\infty$ and thus $T^{1/2}\sigma^2\rho^k/(1-\rho^2) \to 0$.

Then the test statistic is obtained by studentizing S_k.

$$\hat{S} = \left(\hat{\Gamma}_0 \left(\sum_{i=1}^{N} \tilde{e}_{i,t}\tilde{e}_{i,t-k}\right) + 2\sum_{j=1}^{l}(1-j/l)\hat{\Gamma}_j \left(\sum_{i=1}^{N} \tilde{e}_{i,t}\tilde{e}_{i,t-k}\right)\right)^{-1/2} S_k$$

As stated below this test statistic is asymptotically normal under H_0. However, when N is large relative to T there will often be a small sample bias in this test statistic. As discussed in Harris et al. (2005), this arises due to estimation error from the OLS coefficients. The error due to estimation of β_i is of order $T^{-1/2}$. However, when N is not sufficiently small relative to T the aggregation of the estimation error across $i = 1, \ldots, N$ can produce substantial bias even when T is itself large. Harris et al. (2005) propose using the following bias-corrected version of the numerator

$$\begin{aligned} S_k^* &= T^{-1/2}\sum_{i=1}^{N}\sum_{t=k+1}^{T} \tilde{e}_{i,t}\tilde{e}_{i,t-k} \\ &+ T^{-1/2}\sum_{i=1}^{N} tr\left\{\left(T^{-1}\sum_{t=1}^{T} x_{it}x_{it}'\right)^{-1}\left(\hat{\Gamma}_0(\sum_{i=1}^{N} x_{it}\tilde{e}_{it}) + \sum_{j=1}^{l}(1-j/l)(\hat{\Gamma}_j(x_{it}\tilde{e}_{it}) + \hat{\Gamma}_j(x_{it}\tilde{e}_{it})')\right)\right\} \end{aligned}$$

The bias-correction accounts for the individual regression errors that accumulate when aggregated over the N time series. The correction is an estimate of the expectation of the $O_p(T^{-1/2})$ terms in the expansion of S_k. Clearly when the $\hat{\Gamma}_j$ are consistent estimates of the autocorrelations and presuming that the regressors have finite second moments each summand of the second term is $O_p(1)$ and hence the whole term is $O_p(T^{-1/2})$. Let $\hat{S}^*$ denote the statistic obtained by replacing S_k with S_k^* in $\hat{S}$. This is the bias-corrected test statistic.

Harris et al. (2005) show that the bias-corrected statistic is still asymptotically normal since the bias-correction is $o_p(1)$.

We next state the asymptotic results of Harris et al. (2005).

Theorem 1. *Suppose $k = O(T^{1/2})$ and $l = o(k)$. Then under Assumptions 1 and 2, (i) $\hat{S} \to N(0,1)$ under H_0, (ii) $\hat{S}^* \to N(0,1)$ under H_0, and (iii) both $\hat{S}$ and $\hat{S}^*$ diverge to $+\infty$ under H_1.*

This theorem states that both proposed test statistics are asymptotically normal under the null hypothesis and that both are consistent under the alternative that $\phi_i = 1$ for some i. The asymptotics are obtained as an application of a more general limit theory of processes where the degree of autocorrelation is indexed by the sample size developed by Harris, McCabe and Leybourne (2003).

3.3 Monte Carlo simulations

We replicate the finite sample size simulations of Harris et al. (2005) and in addition perform several additional simulations with non-stationary volatility. First we use the dgp of equations (3.1)-(3.2) with $\beta_i = 0$ for all i and $\varepsilon_{it} = \nu_{it} - \theta_i \nu_{i,t-1}$ where $\boldsymbol{\nu_t} = (\nu_{1t}, \ldots, \nu_{Nt})$ is i.i.d. $N(0, \rho)$ where $\rho_{ij} = E(\nu_{it}\nu_{jt})$ and $\rho_{ii} = 1$. In the simulations we vary the MA parameters θ_i and the AR parameters ϕ_i as well as the sample sizes N and T, but fix $\rho_{ij} = 0$. In a second set of exercises we allow the variance to jump from 1 to σ^2 in the middle of the series. That is, $\varepsilon_{it} = \tilde{\nu}_{it} - \theta_i \tilde{\nu}_{i,t-1}$ where $\tilde{\nu}_{it} = \Sigma_t^{1/2} \nu_{it}$, ν_{it} is i.i.d. $N(0, \rho)$ and $\Sigma_t^{1/2} = \{1 + \mathbb{I}(t > \lfloor .5T \rfloor)(\sigma - 1)\} I_N$. We also run simulations allowing for a deterministic time trend. In particular we set $x_{it} = t$ and $\beta_i = \beta$ for all i in the data generating process.

In Tables 1A-1C we used the test statistic $\hat{S}$ for the panels labeled "no bias correction" and $\hat{S}^*$ for the panels labeled "bias-corrected". The residual was obtained by first removing the mean of each series and then normalizing each series of residuals as described in the text. We chose the order of autocorrelation k according to the rule $k = \lfloor (3 * T)^{1/2} \rfloor$ and we chose the lag truncation parameter l according to $l = \lfloor 12 * (T/100)^{1/4} \rfloor$. These are the same choices used in Harris et al. (2005). First examine Table 1A. The first panel of this table should be identical to Table 1(a) of HLM. Indeed both tables show that with the bias-correction the test has nearly exact size in finite samples across various values of θ_i and ϕ_i when $\rho_{ij} = 0$ for $i \neq j$ and the variance is constant over time. Both tables show that the test is the most under-sized when T is small but N is large, $\theta_i = 0$, and $\phi_i = 0.8$. This is likely because the bias of the uncorrected statistic is quite large in this case and the correction is not sufficient. Also, both tables show that the test is most over-sized when sample sizes are small and $\theta_i = 0$ and ϕ_i is small. It can also be seen in the second panel of Table 1A that the uncorrected tests is severely undersized as shown by HLM (cf. Table 1(e) in that paper).

Now consider Tables 1B and 1C. First note that in the second panel, where the results of the uncorrected test are reported, we see that the size in every case is less than the nominal size of 0.05. This suggests that the bias in the uncorrected test statistic caused by not accounting for heteroskedasticity is not *too* large. Indeed we can conclude that the bias is typically negative by comparing with the second panel of Table 1A. Now examine the first panel of Tables 1B and 1C. Note that the failure to account for heteroskedasticity may bias these tests in two ways because both the uncorrected statistic and the bias-correction may be affected by the heteroskedasticity. For example, consider the three columns where the MA parameter is 0. The size is largest in this case but the second panel shows that there is not much of an upward bias in the uncorrected statistic, if at all. Indeed it is apparent that the failure of the bulk of the bias in the "bias-corrected" statistic is due to the use of the wrong model for the correction. Note also though that this second source of bias may be positive or negative; see for example the fifth column. Overall the failure to account for heteroskedasticity did not lead to large size distortions; in the worst case we find a size of less than twice the nominal size.

Tables 2A-2C present new results concerning the model with a deterministic time trend. The residual was obtained by first estimating a regression on a constant and a linear time trend for each series and then normalizing each series of residuals as described in the text. We again chose the order of autocorrelation k according to the rule $k = \lfloor (3 * T)^{1/2} \rfloor$ and we chose the lag truncation parameter l according to $l = \lfloor 12 * (T/100)^{1/4} \rfloor$. Table 2A reports results for the case where the volatility is in fact constant. Note that the bias-correction goes too far in many of the exercises. The size is distorted by more than 2.5 times the nominal size in some cases. However, comparing the results of the volatility jump exercises in Tables 2B and 2C we find the same pattern as we did when no trend was simulated. Namely, the uncorrected statistic, while still in some cases severely undersized, is never over-sized. And the size of the bias-corrected test is typically only slightly higher than the analogous test in Table 2A.

Tables 3 and 4 report the results of the power simulations. To assess the power of the test, both under constant volatility and time-varying volatility, we simulate the model of equations (3.1)-(3.2) with the same parameters as in the size simulations except that we allow some of the N series to have a unit root, i.e., $\phi_i = 1$. We vary the number of series with $\phi_i = 1$ from $M = 1$ to $M = 30$. We also report results where the correlation between the series is $\rho = 0.5$ and $\rho = 0.9$ in addition to the uncorrelated case.

First observe the results in Table 3A. The power number lines up exactly with those reported by Harris et al. (2005), as they should. Note that the power is lower when there is correlation among the separate series in the panel. Yet the power does not drop off toward zero except in the case where N is large and M is small. Compare the results in Tables 3B and 3C. There is an observable drop in power relative to the case of constant volatility. However, the power loss is not substantial in any case. As a baseline, consider the case where $T = 300$, $N = 3$, $\rho = 0$ and $M = 1$. The power drops from .79 to .74 in the first case where the volatility jumps 300% and .72 in the second case where the volatility jumps 500%.

Recall that we found above that the bias-corrected test is slightly over-sized when the series exhibit time-varying variances. We also found, however, that the size of the unadjusted test remains below nominal size under this departure from the model. This is a useful result only if the test also maintains power in this case. To assess this consider panel (b) in Tables 3A-3C. As expected, the lowest power occurs when N is large relative to T - because this is where the small sample bias is largest - and when M is small. This is true whether or not the model is correctly specified, i.e., regardless of the volatility process. However, for large enough T the unadjusted test exhibits substantial power when the series exhibit unattended jumps in the volatility process. When $T = 300$ and there is a 500/% jump in the variance the unadjusted test retains power function which is at modest levels when $M = 1$ and increases quickly in M.

Also, Harris et al. (2005) note that the power results demonstrate clearly the advantage of using the panel test. Observe in Table 3A, panel (a), that when $N = M = 3$ and $T = 75$ the power is 0.69. Increasing N but keeping $M = N$ and $T = 75$, the power increases quickly to 1. This shows that as more evidence of nonstationarity in a panel is added the test is able to reject at a higher rate. Importantly, this result still holds for both the unadjusted and bias-corrected tests when jumps in volatility are present. For example, in Table 3C, panel (a), we see that the power goes from 0.63 to .99 along the same diagonal. In panel (b) of the same table we see that the power is negligible when $t = 75$ but when $T = 150$ the power goes from .72 to 1 along the $N = M$ diagonal.

Table 4 reports power results for the model with a linear time trend. One important message here is that for small T the test has low power, regardless of the volatility process and regardless of whether the bias correction is implemented. Considering then only the panels where $T = 300$, we see that the results mostly mirror those from the case without a trend. There is a small but not substantial power loss due to the jump in volatility in the bias corrected test. The unadjusted test still has non-negligible power particularly as M increases. And the result that when correlation is low the panel test improves power still holds under jumps in volatility.

3.4 Conclusion

Harris et al. (2005) have shown that if the cross-sectional dependence is properly accounted for panel time series data can provide increased power in detecting unit roots. One potential drawback of their procedure, however, is that it fails to account for the possibility of non-stationary volatility in one or more of the observed time series. This is potentially a serious problem because recent work has found that time-varying unconditional variances are quite prevalent in many macroeconomic time series and because other test statistics have been found to exhibit substantial size distortions when this is not accounted for. Through several Monte Carlo simulation exercises we are able to show that panel stationarity tests based on the properly studentized aggregate lagged autocorrelations of multiple series, such as the one proposed by Harris et al. (2005), do not suffer from this same problem. The Harris et al. (2005) test is surprisingly robust without correcting the variance estimator for heteroskedasticity, as usually one would need a HAC (Hetroscedastic and Auocorrelation Consistent) variance estimator to achieve asymptotic robustness.

While we find that the test proposed by Harris et al. (2005) that corrects for small sample bias when N is large relative to T is over-sized, the distortion is minimal. On the other hand, the unadjusted statistic, however, controls the size in all of our simulations. Therefore in some cases it may be preferable to use the unadjusted statistic as a conservative approach to testing for stationarity when time-varying volatility may be present.

One might suspect that though the test controls size despite these departures from the model it may suffer in terms of power when there is time-varying volatility. Our simulation results show that this is in general not the case. While there are slight power losses due to the jumps in volatility the only substantial losses are when $T = 75$. For larger sample sizes the reduction in power is typically smaller, though not negligible. Using the unadjusted test statistic as a conservative strategy when time-varying volatility may be present results in further power losses but in general the power of the test is still substantial and increases when the percentage of nonstationary series in the panel increases.

This certainly raises the question of why the test statistics considered exhibit this robustness property. It would be useful to derive the asymptotic distribution of the statistics under more general assumptions that allow for non-stationary volatility in the residual process. In addition, the simulation results are limited to a particular volatility process with a single break in the middle of the sample. While in some sense this is a particularly difficult volatility process because it is discontinuous, it may be the case that the robustness properties that we find do not generalize. A more extensive simulation study may be warranted. We leave these important extensions for future work.

Table 1A. Empirical size of S(hat) at asymptotic 0.05-level critical values.

Constant Only (MC=10,000)
sigma1/sigma0 = 1 (Constant Volatility)

(a) ρij = 0.0, Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.05	0.06	0.07	0.04	0.04	0.05
5	75	0.05	0.06	0.07	0.04	0.04	0.05
10	75	0.06	0.07	0.06	0.05	0.05	0.06
20	75	0.07	0.07	0.04	0.05	0.05	0.07
30	75	0.07	0.07	0.03	0.05	0.04	0.07
3	150	0.05	0.06	0.06	0.05	0.05	0.05
5	150	0.05	0.06	0.06	0.05	0.04	0.05
10	150	0.05	0.06	0.05	0.05	0.05	0.06
20	150	0.06	0.07	0.05	0.05	0.05	0.06
30	150	0.06	0.07	0.04	0.05	0.05	0.06
3	300	0.05	0.05	0.05	0.05	0.05	0.05
5	300	0.05	0.05	0.06	0.05	0.05	0.05
10	300	0.05	0.05	0.05	0.05	0.05	0.05
20	300	0.05	0.06	0.05	0.05	0.05	0.05
30	300	0.06	0.06	0.04	0.05	0.05	0.05

(e) ρij = 0.0, No Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.02	0.01	0.01	0.03	0.03	0.02
5	75	0.02	0.01	0.00	0.03	0.04	0.02
10	75	0.02	0.01	0.00	0.03	0.04	0.01
20	75	0.01	0.00	0.00	0.03	0.04	0.01
30	75	0.01	0.00	0.00	0.02	0.03	0.01
3	150	0.03	0.03	0.02	0.04	0.05	0.03
5	150	0.03	0.02	0.01	0.04	0.04	0.03
10	150	0.03	0.02	0.00	0.04	0.04	0.02
20	150	0.02	0.01	0.00	0.04	0.05	0.02
30	150	0.02	0.01	0.00	0.04	0.05	0.01
3	300	0.04	0.03	0.03	0.05	0.05	0.04
5	300	0.04	0.03	0.02	0.05	0.05	0.04
10	300	0.03	0.02	0.01	0.05	0.05	0.03
20	300	0.03	0.02	0.00	0.04	0.05	0.02
30	300	0.02	0.01	0.00	0.04	0.05	0.02

Table 1B. Empirical size of S(hat) at asymptotic 0.05-level critical values.

Constant Only (MC=10,000)

sigma1/sigma0 = 3, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.03	0.05	0.09	0.02	0.01	0.04
5	75	0.04	0.07	0.08	0.02	0.02	0.05
10	75	0.06	0.09	0.07	0.03	0.02	0.06
20	75	0.07	0.09	0.04	0.03	0.02	0.08
30	75	0.08	0.09	0.02	0.04	0.02	0.09
3	150	0.04	0.05	0.07	0.04	0.04	0.04
5	150	0.04	0.06	0.07	0.04	0.03	0.05
10	150	0.05	0.07	0.08	0.04	0.03	0.05
20	150	0.06	0.08	0.06	0.04	0.04	0.06
30	150	0.07	0.08	0.06	0.05	0.04	0.07
3	300	0.05	0.05	0.06	0.05	0.05	0.05
5	300	0.05	0.05	0.06	0.05	0.05	0.05
10	300	0.05	0.06	0.07	0.05	0.05	0.05
20	300	0.05	0.06	0.06	0.05	0.05	0.06
30	300	0.06	0.07	0.06	0.05	0.05	0.06

(e) ρij = 0.0, No Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.01	0.01	0.01	0.01	0.01	0.01
5	75	0.01	0.01	0.01	0.01	0.01	0.01
10	75	0.01	0.01	0.00	0.01	0.01	0.01
20	75	0.01	0.00	0.00	0.01	0.01	0.01
30	75	0.01	0.00	0.00	0.01	0.01	0.00
3	150	0.02	0.02	0.02	0.03	0.03	0.03
5	150	0.02	0.02	0.02	0.03	0.03	0.03
10	150	0.02	0.02	0.01	0.03	0.03	0.02
20	150	0.02	0.01	0.00	0.03	0.03	0.02
30	150	0.02	0.01	0.00	0.03	0.03	0.02
3	300	0.04	0.03	0.03	0.04	0.05	0.04
5	300	0.04	0.03	0.03	0.04	0.05	0.04
10	300	0.03	0.03	0.02	0.04	0.04	0.03
20	300	0.03	0.02	0.01	0.04	0.05	0.03
30	300	0.03	0.02	0.01	0.04	0.04	0.02

Table 1C. Empirical size of S(hat) at asymptotic 0.05-level critical values.

Constant Only (MC=10,000)

sigma1/sigma0 = 5, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.03	0.05	0.09	0.01	0.01	0.03
5	75	0.04	0.07	0.08	0.01	0.01	0.05
10	75	0.06	0.09	0.07	0.02	0.01	0.07
20	75	0.07	0.10	0.04	0.02	0.01	0.09
30	75	0.09	0.09	0.02	0.03	0.01	0.09
3	150	0.04	0.05	0.08	0.03	0.03	0.04
5	150	0.04	0.06	0.08	0.04	0.03	0.05
10	150	0.05	0.07	0.08	0.03	0.03	0.05
20	150	0.06	0.08	0.07	0.04	0.03	0.07
30	150	0.07	0.08	0.06	0.04	0.03	0.07
3	300	0.05	0.05	0.07	0.05	0.05	0.05
5	300	0.05	0.05	0.06	0.05	0.05	0.05
10	300	0.05	0.06	0.06	0.05	0.05	0.05
20	300	0.06	0.06	0.06	0.05	0.05	0.06
30	300	0.06	0.07	0.06	0.05	0.05	0.06

(e) ρij = 0.0, No Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.00	0.01	0.01	0.00	0.01	0.01
5	75	0.00	0.01	0.01	0.00	0.00	0.00
10	75	0.00	0.01	0.00	0.00	0.01	0.01
20	75	0.00	0.00	0.00	0.01	0.01	0.01
30	75	0.00	0.00	0.00	0.00	0.00	0.00
3	150	0.02	0.02	0.02	0.03	0.03	0.02
5	150	0.02	0.02	0.02	0.03	0.03	0.02
10	150	0.02	0.02	0.01	0.02	0.03	0.02
20	150	0.02	0.01	0.00	0.03	0.03	0.02
30	150	0.02	0.01	0.00	0.03	0.03	0.02
3	300	0.04	0.03	0.03	0.04	0.05	0.04
5	300	0.04	0.03	0.03	0.05	0.04	0.04
10	300	0.03	0.03	0.02	0.04	0.04	0.03
20	300	0.03	0.02	0.01	0.04	0.04	0.03
30	300	0.03	0.02	0.01	0.04	0.05	0.02

Table 2A. Empirical size of S(hat) at asymptotic 0.05-level critical values.

Linear Trend Included (MC=10,000)
sigma1/sigma0 = 1 (Constant Volatility)

(a) ρij = 0.0, Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.09	0.09	0.04	0.07	0.07	0.08
5	75	0.09	0.09	0.03	0.08	0.08	0.09
10	75	0.11	0.10	0.01	0.09	0.08	0.09
20	75	0.12	0.09	0.00	0.10	0.09	0.09
30	75	0.12	0.08	0.00	0.11	0.09	0.09
3	150	0.07	0.08	0.05	0.07	0.07	0.07
5	150	0.07	0.08	0.04	0.07	0.07	0.07
10	150	0.07	0.08	0.03	0.07	0.07	0.07
20	150	0.09	0.09	0.01	0.08	0.07	0.08
30	150	0.10	0.08	0.01	0.08	0.07	0.08
3	300	0.06	0.07	0.05	0.06	0.06	0.06
5	300	0.06	0.07	0.05	0.06	0.06	0.06
10	300	0.07	0.07	0.04	0.06	0.06	0.06
20	300	0.07	0.07	0.03	0.07	0.07	0.06
30	300	0.07	0.08	0.02	0.06	0.06	0.07

(e) ρij = 0.0, No Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.02	0.01	0.00	0.03	0.03	0.01
5	75	0.01	0.01	0.00	0.03	0.04	0.01
10	75	0.01	0.00	0.00	0.03	0.04	0.01
20	75	0.01	0.00	0.00	0.02	0.04	0.00
30	75	0.00	0.00	0.00	0.02	0.03	0.00
3	150	0.03	0.02	0.00	0.04	0.05	0.02
5	150	0.02	0.01	0.00	0.04	0.04	0.02
10	150	0.02	0.01	0.00	0.03	0.04	0.01
20	150	0.01	0.00	0.00	0.03	0.05	0.01
30	150	0.01	0.00	0.00	0.03	0.04	0.00
3	300	0.03	0.02	0.01	0.04	0.05	0.03
5	300	0.03	0.02	0.00	0.04	0.05	0.03
10	300	0.02	0.01	0.00	0.04	0.05	0.02
20	300	0.02	0.01	0.00	0.04	0.05	0.01
30	300	0.01	0.00	0.00	0.03	0.05	0.01

Table 2B. Empirical size of S(hat) at asymptotic 0.05-level critical values.

Linear Trend Included (MC=10,000)

sigma1/sigma0 = 3, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.08	0.10	0.06	0.06	0.05	0.08
5	75	0.09	0.12	0.05	0.07	0.05	0.09
10	75	0.11	0.12	0.03	0.08	0.06	0.10
20	75	0.13	0.12	0.01	0.09	0.07	0.12
30	75	0.14	0.10	0.00	0.10	0.07	0.11
3	150	0.06	0.08	0.06	0.06	0.06	0.07
5	150	0.07	0.09	0.06	0.06	0.05	0.08
10	150	0.08	0.09	0.04	0.06	0.05	0.08
20	150	0.10	0.10	0.03	0.08	0.06	0.09
30	150	0.10	0.10	0.02	0.08	0.07	0.10
3	300	0.06	0.07	0.06	0.06	0.06	0.06
5	300	0.07	0.07	0.06	0.06	0.06	0.07
10	300	0.07	0.07	0.05	0.06	0.06	0.07
20	300	0.07	0.08	0.04	0.07	0.06	0.07
30	300	0.08	0.09	0.04	0.07	0.06	0.07

(e) ρij = 0.0, No Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.00	0.00	0.00	0.01	0.01	0.00
5	75	0.00	0.00	0.00	0.01	0.01	0.00
10	75	0.00	0.00	0.00	0.01	0.01	0.00
20	75	0.00	0.00	0.00	0.01	0.01	0.00
30	75	0.00	0.00	0.00	0.01	0.01	0.00
3	150	0.02	0.01	0.00	0.03	0.03	0.02
5	150	0.02	0.01	0.00	0.03	0.03	0.02
10	150	0.01	0.01	0.00	0.02	0.03	0.01
20	150	0.01	0.00	0.00	0.02	0.03	0.01
30	150	0.01	0.00	0.00	0.02	0.03	0.00
3	300	0.03	0.02	0.01	0.04	0.05	0.03
5	300	0.03	0.02	0.01	0.04	0.05	0.03
10	300	0.02	0.01	0.00	0.04	0.04	0.02
20	300	0.02	0.01	0.00	0.04	0.05	0.01
30	300	0.02	0.01	0.00	0.03	0.04	0.01

Table 2C. Empirical size of S(hat) at asymptotic 0.05-level critical values.

Linear Trend Included (MC=10,000)

sigma1/sigma0 = 5, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.08	0.10	0.07	0.05	0.04	0.08
5	75	0.09	0.12	0.06	0.06	0.04	0.09
10	75	0.12	0.13	0.03	0.07	0.05	0.11
20	75	0.13	0.12	0.01	0.09	0.06	0.12
30	75	0.15	0.11	0.00	0.10	0.06	0.12
3	150	0.06	0.08	0.07	0.06	0.05	0.07
5	150	0.07	0.09	0.06	0.06	0.05	0.08
10	150	0.08	0.10	0.05	0.06	0.05	0.08
20	150	0.10	0.10	0.03	0.07	0.06	0.09
30	150	0.11	0.10	0.02	0.08	0.06	0.10
3	300	0.06	0.07	0.06	0.06	0.06	0.06
5	300	0.06	0.07	0.06	0.06	0.06	0.07
10	300	0.07	0.07	0.06	0.06	0.06	0.07
20	300	0.07	0.08	0.05	0.06	0.06	0.07
30	300	0.08	0.09	0.04	0.07	0.06	0.07

N	T	ϕi = 0.0 θi = 0.0	ϕi = 0.4 θi = 0.0	ϕi = 0.8 θi = 0.0	ϕi = 0.0 θi = 0.4	ϕi = 0.0 θi = 0.8	ϕi = U[0, 0.8] θi = U[0, 0.8]
3	75	0.00	0.00	0.00	0.00	0.01	0.00
5	75	0.00	0.00	0.00	0.00	0.00	0.00
10	75	0.00	0.00	0.00	0.00	0.01	0.00
20	75	0.00	0.00	0.00	0.00	0.00	0.00
30	75	0.00	0.00	0.00	0.00	0.00	0.00
3	150	0.02	0.01	0.00	0.02	0.03	0.02
5	150	0.02	0.01	0.00	0.02	0.03	0.01
10	150	0.01	0.01	0.00	0.02	0.03	0.01
20	150	0.01	0.00	0.00	0.02	0.03	0.01
30	150	0.01	0.00	0.00	0.02	0.03	0.00
3	300	0.03	0.03	0.01	0.04	0.05	0.03
5	300	0.03	0.02	0.01	0.04	0.04	0.03
10	300	0.02	0.02	0.00	0.04	0.04	0.02
20	300	0.02	0.01	0.00	0.04	0.04	0.02
30	300	0.02	0.01	0.00	0.03	0.04	0.01

Table 3A. Empirical power of S(hat) at asymptotic 0.05-level critical values.

Constant Only (MC=10,000)

sigma1/sigma0 = 1 (Constant Volatility)

(a) ρij = 0.0, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.34	0.69	N/A	N/A	N/A	N/A
5	75	0.28	0.63	0.82	N/A	N/A	N/A
10	75	0.22	0.53	0.75	0.95	N/A	N/A
20	75	0.18	0.42	0.64	0.91	1.00	N/A
30	75	0.17	0.36	0.59	0.85	0.99	1.00
3	150	0.57	0.91	N/A	N/A	N/A	N/A
5	150	0.50	0.89	0.98	N/A	N/A	N/A
10	150	0.40	0.85	0.97	1.00	N/A	N/A
20	150	0.30	0.77	0.94	1.00	1.00	N/A
30	150	0.22	0.68	0.91	1.00	1.00	1.00
3	300	0.79	0.99	N/A	N/A	N/A	N/A
5	300	0.76	0.99	1.00	N/A	N/A	N/A
10	300	0.66	0.98	1.00	1.00	N/A	N/A
20	300	0.53	0.97	1.00	1.00	1.00	N/A
30	300	0.45	0.94	1.00	1.00	1.00	1.00

(b) ρij = 0.0, No Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.10	0.22	N/A	N/A	N/A	N/A
5	75	0.07	0.17	0.26	N/A	N/A	N/A
10	75	0.05	0.11	0.18	0.34	N/A	N/A
20	75	0.03	0.06	0.11	0.24	0.45	N/A
30	75	0.02	0.05	0.09	0.19	0.36	0.52
3	150	0.40	0.78	N/A	N/A	N/A	N/A
5	150	0.32	0.74	0.90	N/A	N/A	N/A
10	150	0.22	0.64	0.86	0.99	N/A	N/A
20	150	0.13	0.48	0.76	0.98	1.00	N/A
30	150	0.08	0.37	0.68	0.95	1.00	1.00
3	300	0.73	0.98	N/A	N/A	N/A	N/A
5	300	0.67	0.97	1.00	N/A	N/A	N/A
10	300	0.55	0.96	1.00	1.00	N/A	N/A
20	300	0.39	0.92	0.99	1.00	1.00	N/A
30	300	0.27	0.88	0.99	1.00	1.00	1.00

Table 3A, cont'd.

(c) ρij = 0.5, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.31	0.62	N/A	N/A	N/A	N/A
5	75	0.22	0.55	0.68	N/A	N/A	N/A
10	75	0.14	0.39	0.58	0.75	N/A	N/A
20	75	0.09	0.21	0.35	0.65	0.82	N/A
30	75	0.07	0.14	0.23	0.52	0.76	0.84
3	150	0.54	0.87	N/A	N/A	N/A	N/A
5	150	0.41	0.82	0.93	N/A	N/A	N/A
10	150	0.23	0.70	0.89	0.97	N/A	N/A
20	150	0.13	0.41	0.67	0.93	0.99	N/A
30	150	0.09	0.27	0.48	0.85	0.98	0.99
3	300	0.77	0.98	N/A	N/A	N/A	N/A
5	300	0.69	0.98	1.00	N/A	N/A	N/A
10	300	0.43	0.94	0.99	1.00	N/A	N/A
20	300	0.21	0.75	0.95	1.00	1.00	N/A
30	300	0.14	0.52	0.84	0.99	1.00	1.00

(d) ρij = 0.9, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.30	0.50	N/A	N/A	N/A	N/A
5	75	0.17	0.44	0.48	N/A	N/A	N/A
10	75	0.09	0.25	0.40	0.49	N/A	N/A
20	75	0.07	0.12	0.21	0.41	0.50	N/A
30	75	0.05	0.08	0.13	0.28	0.48	0.50
3	150	0.50	0.70	N/A	N/A	N/A	N/A
5	150	0.32	0.69	0.71	N/A	N/A	N/A
10	150	0.17	0.50	0.64	0.71	N/A	N/A
20	150	0.08	0.21	0.44	0.67	0.75	N/A
30	150	0.07	0.14	0.26	0.56	0.71	0.75
3	300	0.76	0.90	N/A	N/A	N/A	N/A
5	300	0.56	0.88	0.92	N/A	N/A	N/A
10	300	0.28	0.76	0.89	0.93	N/A	N/A
20	300	0.13	0.46	0.72	0.89	0.93	N/A
30	300	0.10	0.28	0.52	0.81	0.92	0.93

Table 3B. Empirical power of S(hat) at asymptotic 0.05-level critical values.

Constant Only (MC=10,000)

sigma1/sigma0 = 3, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	M = 1	M = 3	M = 5	M = 10	M = 20	M = 30
3	75	0.33	0.65	N/A	N/A	N/A	N/A
5	75	0.28	0.57	0.75	N/A	N/A	N/A
10	75	0.23	0.48	0.66	0.89	N/A	N/A
20	75	0.19	0.39	0.55	0.81	0.97	N/A
30	75	0.19	0.34	0.47	0.76	0.95	1.00
3	150	0.55	0.88	N/A	N/A	N/A	N/A
5	150	0.46	0.85	0.96	N/A	N/A	N/A
10	150	0.36	0.79	0.94	1.00		
20	150	0.27	0.68	0.89	0.99	1.00	N/A
30	150	0.23	0.60	0.83	0.99	1.00	1.00
3	300	0.74	0.98	N/A	N/A	N/A	N/A
5	300	0.68	0.97	1.00	N/A	N/A	N/A
10	300	0.58	0.96	1.00	1.00	N/A	N/A
20	300	0.45	0.93	0.99	1.00	1.00	N/A
30	300	0.37	0.87	0.99	1.00	1.00	1.00

(b) ρij = 0.0, No Bias Correction

N	T	M = 1	M = 3	M = 5	M = 10	M = 20	M = 30
3	75	0.10	0.17	N/A	N/A	N/A	N/A
5	75	0.08	0.14	0.17	N/A	N/A	N/A
10	75	0.05	0.11	0.14	0.15	N/A	N/A
20	75	0.03	0.07	0.10	0.13	0.13	N/A
30	75	0.02	0.06	0.08	0.10	0.10	0.08
3	150	0.38	0.75	N/A	N/A	N/A	N/A
5	150	0.30	0.69	0.86	N/A	N/A	N/A
10	150	0.20	0.57	0.80	0.97	N/A	N/A
20	150	0.12	0.42	0.66	0.94	1.00	N/A
30	150	0.08	0.32	0.57	0.90	1.00	1.00
3	300	0.67	0.96	N/A	N/A	N/A	N/A
5	300	0.61	0.95	0.99	N/A	N/A	N/A
10	300	0.47	0.92	0.99	1.00	N/A	N/A
20	300	0.31	0.85	0.98	1.00	1.00	N/A
30	300	0.22	0.79	0.96	1.00	1.00	1.00

Table 3B, cont'd.

(c) ρij = 0.5, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.30	0.59	N/A	N/A	N/A	N/A
5	75	0.25	0.52	0.66	N/A	N/A	N/A
10	75	0.14	0.38	0.53	0.70	N/A	N/A
20	75	0.09	0.21	0.34	0.59	0.77	N/A
30	75	0.07	0.15	0.25	0.48	0.71	0.78
3	150	0.50	0.83	N/A	N/A	N/A	N/A
5	150	0.39	0.79	0.89	N/A	N/A	N/A
10	150	0.22	0.64	0.82	0.94	N/A	N/A
20	150	0.11	0.38	0.61	0.90	0.98	N/A
30	150	0.09	0.24	0.41	0.80	0.95	0.98
3	300	0.71	0.96	N/A	N/A	N/A	N/A
5	300	0.61	0.94	0.99	N/A	N/A	N/A
10	300	0.37	0.89	0.97	1.00	N/A	N/A
20	300	0.19	0.62	0.89	1.00	1.00	N/A
30	300	0.12	0.44	0.75	0.97	1.00	1.00

(d) ρij = 0.9, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.28	0.51	N/A	N/A	N/A	N/A
5	75	0.17	0.42	0.52	N/A	N/A	N/A
10	75	0.09	0.28	0.40	0.53	N/A	N/A
20	75	0.05	0.12	0.23	0.41	0.51	N/A
30	75	0.03	0.08	0.14	0.30	0.48	0.49
3	150	0.48	0.69	N/A	N/A	N/A	N/A
5	150	0.29	0.64	0.70	N/A	N/A	N/A
10	150	0.12	0.46	0.62	0.70	N/A	N/A
20	150	0.07	0.22	0.40	0.61	0.72	N/A
30	150	0.06	0.13	0.25	0.51	0.70	0.72
3	300	0.68	0.84	N/A	N/A	N/A	N/A
5	300	0.50	0.83	0.88	N/A	N/A	N/A
10	300	0.22	0.69	0.83	0.89	N/A	N/A
20	300	0.12	0.37	0.62	0.84	0.89	N/A
30	300	0.09	0.22	0.43	0.75	0.88	0.89

Table 3C. Empirical power of S(hat) at asymptotic 0.05-level critical values.

Constant Only (MC=10,000)

sigma1/sigma0 = 5, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	M = 1	3	5	10	20	30
3	75	0.33	0.63	N/A	N/A	N/A	N/A
5	75	0.28	0.56	0.73	N/A	N/A	N/A
10	75	0.22	0.46	0.63	0.86	N/A	N/A
20	75	0.19	0.37	0.52	0.77	0.95	N/A
30	75	0.17	0.32	0.44	0.68	0.92	0.99
3	150	0.54	0.87	N/A	N/A	N/A	N/A
5	150	0.45	0.84	0.95	N/A	N/A	N/A
10	150	0.35	0.76	0.92	1.00	N/A	N/A
20	150	0.26	0.66	0.87	0.99	1.00	N/A
30	150	0.22	0.60	0.81	0.98	1.00	1.00
3	300	0.72	0.97	N/A	N/A	N/A	N/A
5	300	0.66	0.96	1.00	N/A	N/A	N/A
10	300	0.56	0.94	0.99	1.00	N/A	N/A
20	300	0.43	0.90	0.99	1.00	1.00	N/A
30	300	0.35	0.86	0.98	1.00	1.00	1.00

(b) ρij = 0.0, No Bias Correction

N	T	M = 1	3	5	10	20	30
3	75	0.10	0.14	N/A	N/A	N/A	N/A
5	75	0.08	0.13	0.12	N/A	N/A	N/A
10	75	0.05	0.10	0.11	0.10	N/A	N/A
20	75	0.03	0.07	0.10	0.09	0.06	N/A
30	75	0.02	0.05	0.07	0.09	0.07	0.03
3	150	0.37	0.72	N/A	N/A	N/A	N/A
5	150	0.30	0.67	0.85	N/A	N/A	N/A
10	150	0.20	0.55	0.77	0.96	N/A	N/A
20	150	0.11	0.40	0.63	0.92	1.00	N/A
30	150	0.08	0.33	0.53	0.86	0.99	1.00
3	300	0.64	0.94	N/A	N/A	N/A	N/A
5	300	0.57	0.93	0.99	N/A	N/A	N/A
10	300	0.46	0.90	0.98	1.00	N/A	N/A
20	300	0.30	0.83	0.96	1.00	1.00	N/A
30	300	0.23	0.74	0.95	1.00	1.00	1.00

Table 3C, cont'd.

(c) ρij = 0.5, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.32	0.57	N/A	N/A	N/A	N/A
5	75	0.24	0.52	0.63	N/A	N/A	N/A
10	75	0.15	0.38	0.54	0.71	N/A	N/A
20	75	0.09	0.21	0.35	0.56	0.74	N/A
30	75	0.07	0.16	0.27	0.44	0.67	0.74
3	150	0.51	0.81	N/A	N/A	N/A	N/A
5	150	0.37	0.77	0.88	N/A	N/A	N/A
10	150	0.22	0.61	0.81	0.94	N/A	N/A
20	150	0.13	0.36	0.59	0.89	0.97	N/A
30	150	0.09	0.24	0.41	0.74	0.95	0.97
3	300	0.70	0.95	N/A	N/A	N/A	N/A
5	300	0.58	0.93	0.98	N/A	N/A	N/A
10	300	0.35	0.86	0.96	1.00	N/A	N/A
20	300	0.19	0.58	0.87	0.99	1.00	N/A
30	300	0.13	0.41	0.68	0.96	1.00	1.00

(d) ρij = 0.9, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.28	0.49	N/A	N/A	N/A	N/A
5	75	0.16	0.42	0.48	N/A	N/A	N/A
10	75	0.08	0.27	0.40	0.51	N/A	N/A
20	75	0.04	0.13	0.24	0.39	0.50	N/A
30	75	0.04	0.07	0.13	0.28	0.46	0.50
3	150	0.44	0.65	N/A	N/A	N/A	N/A
5	150	0.29	0.64	0.67	N/A	N/A	N/A
10	150	0.14	0.45	0.60	0.70	N/A	N/A
20	150	0.07	0.22	0.34	0.60	0.69	N/A
30	150	0.06	0.13	0.23	0.47	0.66	0.71
3	300	0.68	0.84	N/A	N/A	N/A	N/A
5	300	0.47	0.80	0.85	N/A	N/A	N/A
10	300	0.22	0.66	0.81	0.87	N/A	N/A
20	300	0.11	0.35	0.61	0.81	0.88	N/A
30	300	0.09	0.22	0.38	0.72	0.87	0.87

Table 4A. Empirical power of S(hat) at asymptotic 0.05-level critical values.

Linear Trend Included (MC=10,000)

sigma1/sigma0 = 1 (Constant Volatility)

(a) ρij = 0.0, Bias Correction

N	T	M: 1	3	5	10	20	30
3	75	0.09	0.06	N/A	N/A	N/A	N/A
5	75	0.09	0.07	0.04	N/A	N/A	N/A
10	75	0.11	0.08	0.06	0.02	N/A	N/A
20	75	0.11	0.09	0.07	0.03	0.00	N/A
30	75	0.12	0.10	0.09	0.04	0.00	0.00
3	150	0.24	0.42	N/A	N/A	N/A	N/A
5	150	0.22	0.40	0.51	N/A	N/A	N/A
10	150	0.18	0.35	0.47	0.66	N/A	N/A
20	150	0.16	0.30	0.41	0.63	0.83	N/A
30	150	0.14	0.27	0.38	0.57	0.80	0.91
3	300	0.52	0.84	N/A	N/A	N/A	N/A
5	300	0.47	0.82	0.94	N/A	N/A	N/A
10	300	0.39	0.78	0.92	0.99	N/A	N/A
20	300	0.31	0.71	0.89	0.99	1.00	N/A
30	300	0.26	0.65	0.86	0.99	1.00	1.00

(b) ρij = 0.0, No Bias Correction

N	T	M: 1	3	5	10	20	30
3	75	0.00	0.00	N/A	N/A	N/A	N/A
5	75	0.00	0.00	0.00	N/A	N/A	N/A
10	75	0.00	0.00	0.00	0.00	N/A	N/A
20	75	0.00	0.00	0.00	0.00	0.00	N/A
30	75	0.00	0.00	0.00	0.00	0.00	0.00
3	150	0.03	0.03	N/A	N/A	N/A	N/A
5	150	0.03	0.03	0.02	N/A	N/A	N/A
10	150	0.03	0.02	0.02	0.01	N/A	N/A
20	150	0.02	0.01	0.01	0.01	0.00	N/A
30	150	0.01	0.01	0.01	0.01	0.00	0.00
3	300	0.30	0.57	N/A	N/A	N/A	N/A
5	300	0.25	0.53	0.70	N/A	N/A	N/A
10	300	0.17	0.45	0.65	0.89	N/A	N/A
20	300	0.10	0.33	0.53	0.84	0.98	N/A
30	300	0.07	0.24	0.44	0.80	0.98	1.00

Table 4A, cont'd.

(c) ρij = 0.5, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.08	0.07	N/A	N/A	N/A	N/A
5	75	0.09	0.08	0.07	N/A	N/A	N/A
10	75	0.09	0.10	0.08	0.06	N/A	N/A
20	75	0.10	0.12	0.11	0.06	0.05	N/A
30	75	0.10	0.10	0.11	0.10	0.07	0.05
3	150	0.23	0.36	N/A	N/A	N/A	N/A
5	150	0.18	0.33	0.40	N/A	N/A	N/A
10	150	0.14	0.27	0.35	0.46	N/A	N/A
20	150	0.10	0.18	0.27	0.41	0.48	N/A
30	150	0.10	0.14	0.21	0.33	0.43	0.51
3	300	0.49	0.76	N/A	N/A	N/A	N/A
5	300	0.41	0.75	0.81	N/A	N/A	N/A
10	300	0.26	0.63	0.80	0.89	N/A	N/A
20	300	0.15	0.42	0.64	0.87	0.94	N/A
30	300	0.12	0.30	0.49	0.78	0.92	0.95

(d) ρij = 0.9, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.10	0.09	N/A	N/A	N/A	N/A
5	75	0.11	0.08	0.09	N/A	N/A	N/A
10	75	0.10	0.10	0.11	0.08	N/A	N/A
20	75	0.10	0.11	0.10	0.10	0.08	N/A
30	75	0.08	0.09	0.10	0.09	0.09	0.08
3	150	0.23	0.29	N/A	N/A	N/A	N/A
5	150	0.16	0.29	0.32	N/A	N/A	N/A
10	150	0.12	0.21	0.28	0.31	N/A	N/A
20	150	0.08	0.13	0.19	0.28	0.31	N/A
30	150	0.07	0.10	0.15	0.22	0.30	0.31
3	300	0.47	0.60	N/A	N/A	N/A	N/A
5	300	0.32	0.60	0.61	N/A	N/A	N/A
10	300	0.18	0.46	0.56	0.60	N/A	N/A
20	300	0.10	0.29	0.41	0.57	0.61	N/A
30	300	0.08	0.18	0.31	0.50	0.60	0.63

Table 4B. Empirical power of S(hat) at asymptotic 0.05-level critical values.

Linear Trend Included (MC=10,000)
sigma1/sigma0 = 3, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	M=1	3	5	10	20	30
3	75	0.11	0.08	N/A	N/A	N/A	N/A
5	75	0.10	0.09	0.07	N/A	N/A	N/A
10	75	0.15	0.11	0.07	0.03	N/A	N/A
20	75	0.15	0.13	0.10	0.06	0.01	N/A
30	75	0.14	0.15	0.11	0.06	0.01	0.00
3	150	0.26	0.40	N/A	N/A	N/A	N/A
5	150	0.23	0.39	0.48	N/A	N/A	N/A
10	150	0.19	0.34	0.44	0.62	N/A	N/A
20	150	0.17	0.32	0.39	0.56	0.74	N/A
30	150	0.17	0.29	0.35	0.49	0.70	0.80
3	300	0.49	0.79	N/A	N/A	N/A	N/A
5	300	0.43	0.78	0.91	N/A	N/A	N/A
10	300	0.37	0.71	0.88	0.98	N/A	N/A
20	300	0.28	0.66	0.84	0.98	1.00	N/A
30	300	0.24	0.57	0.79	0.97	1.00	1.00

(b) ρij = 0.0, No Bias Correction

N	T	M=1	3	5	10	20	30
3	75	0.00	0.00	N/A	N/A	N/A	N/A
5	75	0.00	0.00	0.00	N/A	N/A	N/A
10	75	0.00	0.00	0.00	0.00	N/A	N/A
20	75	0.00	0.00	0.00	0.00	0.00	N/A
30	75	0.00	0.00	0.00	0.00	0.00	0.00
3	150	0.04	0.04	N/A	N/A	N/A	N/A
5	150	0.04	0.04	0.03	N/A	N/A	N/A
10	150	0.03	0.03	0.03	0.02	N/A	N/A
20	150	0.02	0.03	0.03	0.02	0.00	N/A
30	150	0.01	0.02	0.02	0.02	0.00	0.00
3	300	0.27	0.53	N/A	N/A	N/A	N/A
5	300	0.23	0.49	0.66	N/A	N/A	N/A
10	300	0.15	0.43	0.60	0.83	N/A	N/A
20	300	0.09	0.29	0.49	0.77	0.96	N/A
30	300	0.06	0.21	0.40	0.70	0.94	0.99

Table 4B, cont'd.

(c) ρij = 0.5, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.12	0.09	N/A	N/A	N/A	N/A
5	75	0.11	0.10	0.08	N/A	N/A	N/A
10	75	0.10	0.12	0.11	0.08	N/A	N/A
20	75	0.10	0.12	0.12	0.11	0.07	N/A
30	75	0.10	0.12	0.13	0.14	0.09	0.06
3	150	0.24	0.38	N/A	N/A	N/A	N/A
5	150	0.20	0.34	0.42	N/A	N/A	N/A
10	150	0.15	0.26	0.37	0.42	N/A	N/A
20	150	0.12	0.18	0.26	0.41	0.47	N/A
30	150	0.09	0.17	0.22	0.35	0.44	0.48
3	300	0.48	0.72	N/A	N/A	N/A	N/A
5	300	0.40	0.69	0.78	N/A	N/A	N/A
10	300	0.22	0.56	0.74	0.86	N/A	N/A
20	300	0.13	0.36	0.56	0.81	0.90	N/A
30	300	0.12	0.27	0.42	0.71	0.89	0.92

(d) ρij = 0.9, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.11	0.10	N/A	N/A	N/A	N/A
5	75	0.10	0.11	0.10	N/A	N/A	N/A
10	75	0.10	0.13	0.12	0.11	N/A	N/A
20	75	0.07	0.11	0.12	0.13	0.11	N/A
30	75	0.07	0.09	0.09	0.12	0.12	0.11
3	150	0.20	0.31	N/A	N/A	N/A	N/A
5	150	0.16	0.29	0.30	N/A	N/A	N/A
10	150	0.12	0.23	0.27	0.32	N/A	N/A
20	150	0.09	0.13	0.21	0.28	0.30	N/A
30	150	0.08	0.11	0.15	0.23	0.29	0.32
3	300	0.42	0.56	N/A	N/A	N/A	N/A
5	300	0.30	0.57	0.59	N/A	N/A	N/A
10	300	0.16	0.45	0.56	0.58	N/A	N/A
20	300	0.11	0.23	0.37	0.55	0.59	N/A
30	300	0.08	0.16	0.24	0.47	0.59	0.60

Table 4C. Empirical power of S(hat) at asymptotic 0.05-level critical values.

Linear Trend Included (MC=10,000)

sigma1/sigma0 = 5, tauV = 0.5

(a) ρij = 0.0, Bias Correction

N	T	M = 1	3	5	10	20	30
3	75	0.12	0.08	N/A	N/A	N/A	N/A
5	75	0.14	0.10	0.08	N/A	N/A	N/A
10	75	0.14	0.12	0.09	0.03	N/A	N/A
20	75	0.15	0.15	0.11	0.05	0.01	N/A
30	75	0.16	0.15	0.12	0.07	0.02	0.00
3	150	0.25	0.40	N/A	N/A	N/A	N/A
5	150	0.20	0.39	0.50	N/A	N/A	N/A
10	150	0.18	0.35	0.44	0.59	N/A	N/A
20	150	0.16	0.28	0.37	0.55	0.70	N/A
30	150	0.17	0.28	0.35	0.48	0.68	0.78
3	300	0.47	0.80	N/A	N/A	N/A	N/A
5	300	0.42	0.75	0.90	N/A	N/A	N/A
10	300	0.36	0.71	0.87	0.98	N/A	N/A
20	300	0.27	0.62	0.81	0.98	1.00	N/A
30	300	0.24	0.56	0.78	0.97	1.00	1.00

(b) ρij = 0.0, No Bias Correction

N	T	M = 1	3	5	10	20	30
3	75	0.00	0.00	N/A	N/A	N/A	N/A
5	75	0.00	0.00	0.00	N/A	N/A	N/A
10	75	0.00	0.00	0.00	0.00	N/A	N/A
20	75	0.00	0.00	0.00	0.00	0.00	N/A
30	75	0.00	0.00	0.00	0.00	0.00	0.00
3	150	0.05	0.05	N/A	N/A	N/A	N/A
5	150	0.04	0.04	0.03	N/A	N/A	N/A
10	150	0.03	0.04	0.04	0.02	N/A	N/A
20	150	0.02	0.03	0.03	0.02	0.01	N/A
30	150	0.02	0.02	0.03	0.02	0.01	0.00
3	300	0.28	0.52	N/A	N/A	N/A	N/A
5	300	0.23	0.48	0.63	N/A	N/A	N/A
10	300	0.15	0.40	0.58	0.81	N/A	N/A
20	300	0.09	0.30	0.46	0.74	0.95	N/A
30	300	0.06	0.22	0.38	0.68	0.93	0.99

Table 4C, cont'd.

(c) ρij = 0.5, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.12	0.10	N/A	N/A	N/A	N/A
5	75	0.11	0.12	0.10	N/A	N/A	N/A
10	75	0.12	0.14	0.13	0.08	N/A	N/A
20	75	0.12	0.13	0.16	0.11	0.09	N/A
30	75	0.11	0.13	0.13	0.14	0.11	0.07
3	150	0.22	0.36	N/A	N/A	N/A	N/A
5	150	0.20	0.35	0.41	N/A	N/A	N/A
10	150	0.14	0.26	0.35	0.42	N/A	N/A
20	150	0.10	0.20	0.27	0.38	0.45	N/A
30	150	0.12	0.14	0.22	0.31	0.44	0.47
3	300	0.46	0.68	N/A	N/A	N/A	N/A
5	300	0.35	0.68	0.78	N/A	N/A	N/A
10	300	0.23	0.57	0.72	0.84	N/A	N/A
20	300	0.13	0.36	0.56	0.80	0.89	N/A
30	300	0.10	0.21	0.41	0.70	0.87	0.92

(d) ρij = 0.9, Bias Correction

N	T	M					
		1	3	5	10	20	30
3	75	0.13	0.12	N/A	N/A	N/A	N/A
5	75	0.12	0.13	0.13	N/A	N/A	N/A
10	75	0.09	0.13	0.13	0.11	N/A	N/A
20	75	0.07	0.11	0.11	0.13	0.12	N/A
30	75	0.07	0.09	0.11	0.12	0.13	0.12
3	150	0.22	0.32	N/A	N/A	N/A	N/A
5	150	0.16	0.28	0.31	N/A	N/A	N/A
10	150	0.11	0.22	0.28	0.31	N/A	N/A
20	150	0.09	0.14	0.19	0.30	0.32	N/A
30	150	0.07	0.11	0.15	0.22	0.31	0.32
3	300	0.44	0.58	N/A	N/A	N/A	N/A
5	300	0.28	0.55	0.58	N/A	N/A	N/A
10	300	0.16	0.40	0.52	0.57	N/A	N/A
20	300	0.09	0.24	0.37	0.53	0.56	N/A
30	300	0.09	0.16	0.26	0.45	0.56	0.60

Chapter 4

Detecting multiple level breaks in the presence of non-stationary volatility

In this paper we analyze the impact of non-stationary volatility on a recently developed procedure for testing the null hypothesis of no break in level against the alternative of (possibly) multiple levels breaks, occurring at unknown point(s) in the sample. The procedure derived by Harvey, Leybourne and Taylor (2010), is a combination of two unit root tests: one designed for $I(0)$ processes and the other for $I(1)$ ones. The procedure takes a *union of rejections* approach whereby the null hypothesis of no level break is rejected if either of the two tests rejects. In its analysis, Harvey et al. (2010) assumes the shocks follow a linear process driven by i.i.d innovations. Using Monte Carlo simulations we show that the tests are oversized when presented with non stationary innovations, specifically in a form of a one-time change in the volatility. As a solution to the inference problem we propose the wild bootstrap implementation of the procedure, using the level break estimator from the original data. The wild bootstrap method does not require the practitioner to specify a parametric model for volatility and is shown to produce good size and power, and performs well in practice.

4.1 Introduction

One must account for structural changes in the parameters of any model if accurate inference is of interest. The first wave of research has been focusing on tests for structural change in the parameters of stationary forms see, *inter alia*, Stock (1994), Kuan and Hornik (1995) and Perron (2006). Stock and Watson (1996, 1999) and Perron and Zhu (2005), *inter alia*, find that many economic and financial data exhibit temporary stationary characteristics or possess permanent unit roots. This motivated a new area of research that aims at deriving tests for structural changes in parameters that are robust to whether the series follows a stationary process or has a unit root. Similar to Models B and C of Perron (1989), Harvey, Leybourne and Taylor (2009a) develop tests for one-time break in the slope of the deterministic trend function.

Harvey et al. (2010), henceforth HLT, extend some of this initial work to allow for multiple level breaks occurring at unknown break points, while maintaining validity for both $I(0)$ and $I(1)$ stochastic processes. It has recently been recognized that series that have breaks in level tend to also exhibit breaks in the unconditional volatility of the shocks driving the process (Busetti and Taylor 2003, Sensier and Dijk 2004). This suggests a possible shortcoming in the test developed by HLT since they assume a non time-varying volatility. Non-constant volatility has been found to produce significant size distortions in standard unit root and stationary tests (Kim et al. 2002, Busetti and Taylor 2003, Cavaliere 2004, Cavaliere and Taylor 2005, Cavaliere and Taylor 2007, Cavaliere and Taylor 2008). Cavaliere, Harvey, Leybourne and Taylor (2011) derive the distribution of the unit root test of Harris, Harvey, Leybourne and

Taylor (2009) when time-varying volatility is present.

In Monte Carlo studies we find similar size distortions in the HLT test when the volatility of the driving process exhibits a break. Similarly to Cavaliere and Taylor (2008) and Cavaliere et al. (2011) we develop a testing procedure based on the wild bootstrap in order to correct for these size distortions.

The remainder of the paper is organized as follows. Section 2 lays out the multiple level model of HLT. In Section 3 we discuss the test statistic proposed by HLT for constant volatility of the driving process. In Section 4 we demonstrate the performance of this test when the volatility is not constant and propose a solution based on two separate wild bootstrap procedures. Section 5 concludes.

In the remainder of the paper $\lfloor \cdot \rfloor$ will denote the integer part and $\stackrel{d}{\to}$ and $\stackrel{p}{\to}$ will denote convergence in distribution and convergence in probability, respectively. We will also use $\mathcal{D}$ to denote the space of all processes on $[0, 1]$ that are right continuous with left limits.

4.2 A multiple level breaks model with non-stationary volatility

We follow the model representation of HLT with the exception that we allow the volatility of the innovations to exhibit non-stationary behaviour. To be specific about the nature of this non-stationary behavior we use the construction of Cavaliere and Taylor (2007), Cavaliere and Taylor (2008), and Cavaliere et al. (2011). In the HLT model, T observations from a time series process

$\{y_t\}$ satisfy

$$y_t = \alpha + \sum_{i=1}^{n} \gamma_{i,T}^{*} DU_t(\lfloor \tau_i T \rfloor) + u_t, t = 1, ..., T \tag{4.1}$$

$$u_t = \rho_T u_{t-1} + \varepsilon_t, t = 2, ..., T \tag{4.2}$$

where $DU_t(\lfloor \tau_i T \rfloor) = 1(t > \lfloor \tau_i T \rfloor)$ with $\lfloor \tau_i T \rfloor$ a potential level break point with fraction τ_i and magnitude $\gamma_{i,T}^{*}$. The level breaks points are unknown but are bounded away from 0 and 1. That is, $\tau_i \in \Lambda = \lfloor \tau_L, \tau_U \rfloor$ where $0 < \tau_L < \tau_U < 1$;. The fractions τ_L and τ_U represent the lower and upper trimming parameters, below and above which no break is assumed to take place. Without loss of generality we can order that break points as follows: $\tau_1 < \tau_2 < \cdots < \tau_{n-1} < \tau_n$. As is typical in the literature we require a condition on initial condition, u_1. Following HLT we assume that $T^{-1/2} u_1 \xrightarrow{p} 0$, so that the first observation does not dominate the stochastic process and becomes negligible as T gets larger. The error process $\{\varepsilon_t\}$ is taken to satisfy the following conventional linear process assumption.

Assumption *LP***:** *The stochastic process* $\{\varepsilon_t\}$ *is such that* $\varepsilon_t = C(L)\eta_t$, where $\eta_t = \sigma_t z_t$ and $C(L) := \sum_{j=0}^{\infty} C_j L^j$ *with* $C(1)^2 > 0$ and $\sum_{i=0}^{\infty} i|C_i| < \infty$, *and where* $\{z_t\}$ *is an i.i.d. sequence with mean zero and unit variance and finite fourth moment.*

Note that under this assumption, η_t has mean 0 and time-varying variance σ_t. We assume this time-varying variance satisfies the following mild condition.

Assumption $\mathcal{A}$: *The volatility term σ_t satisfies $\sigma_t = \omega(t/T)$ where $\omega(\cdot) \in \mathcal{D}$ is non-stochastic and strictly positive. For $t > 0, \sigma_t \leq \tilde{\sigma} < \infty$.*

The conditions on the volatility process given in this assumption are the same as those given in Cavaliere and Taylor (2008) and Cavaliere et al. (2011) (see Assumption $\mathcal{A}_3$ in either paper). Note that the data-generating process assumed in HLT is reproduced here if $\omega(s) = \sigma$. In general, however, this assumption allows for a broader class of models than HLT. The restrictions on the volatility process are quite weak. The assumption only requires that the variance is bounded and exhibits at most a countable number of jumps. See Cavaliere and Taylor (2007) for more discussion of the class of models that satisfy this assumption.

Note that the autoregressive factor in (4.2) depends on the sample size T. This allows us to study the local-to-unit root asymptotics. As in HLT, two cases for the order of integration of the autoregressive process, u_t, are considered that amount to two different assumption on ρ_T.

(a) The $I(1)$ case for u_t is represented by setting $\rho_T := 1-c/T$ for $0 \leq c < \infty$ in (4.2), which permits (local to) unit root behavior when ($c > 0$) $c = 0$. It is also assumed that $\gamma_{i,T}^* := \omega_\varepsilon T^{1/2}\gamma_i$, $i = 1, \ldots, n$. The $T^{1/2}$ scaling in $\gamma_{i,T}^*$ provides the appropriate Pitman drift so that the test statistics grow at the right rate as T gets bigger in order for local power to be assessed, while scaling by ω_ε is a convenient device allowing it to be factored out when the limit distribution for this process is derived.

(b) The $I(0)$ case for u_t is achieved by setting $\rho_T = \rho$ for all T where $|\rho| < 1$ in (4.2), where the long run variance of u_t is given by $\omega_u^2 := \lim_{T\to\infty} T^{-1}E(\sum_{t=1}^{T} u_t)^2 = \sigma_\eta^2 C(1)^2/(1-\rho)^2$. It is assumed that $\gamma_{i,T}^* := \omega_u T^{-1/2}\gamma_i, i = 1, \ldots, n$, with $T^{-1/2}$ now providing the appropriate Pitan drift, and scaling by ω_u again being used for convenience in the derivation of the limit distribution.

Formally, the two cases are embodied in the following assumptions:

Assumption I(1) *Let Assumption LP hold. Also, let* $\rho_T = 1 - c/T$, $0 \leq c < \infty$, *and let* $\gamma_{i,T}^* = \omega_\varepsilon T^{1/2}\gamma_i$.

Assumption I(0) *Let Assumption LP hold. Also, let* $\rho_T = \rho$ *where* $|\rho| < 1$ *and let* $\gamma_{i,T}^* = \omega_u T^{-1/2}\gamma_i$.

In the next section we discuss HLT's proposed statistic for detecting the level breaks in this model. They propose a procedure that is valid under either Assumption I(1) or Assumption I(0). Their procedure is not valid, however, if $\omega(s)$ is not constant. In Section 4 we discuss how to adjust the procedure to account for this new aspect of the model.

4.3 The HLT procedure for testing for multiple level breaks

Testing for level break(s) is carried out by testing the null hypothesis of no level breaks, that is

$$H_0 : \gamma^*_{i,T} = 0,$$

for all $i \in \{1, 2, \ldots, n\}$, against the alternative of at least one level break; that is

$$H_1 : \gamma^*_{i,T} \neq 0$$

for at least one $i \in \{1, \ldots, n\}$

In implementing a test of such hypothesis, HLT consider a sequence of statistics that belong to the *generalized fluctuations* class of test statistics for structural change, introduced in Kuan and Hornik (1995) and Leisch, Hornik and Kuan (2000), *inter alia*. The statistic takes the form of $M_{t,\lfloor mT \rfloor}$, for $t \in \Lambda_T := [\lfloor \tau_L T \rfloor, \lfloor \tau_U T \rfloor]$, where

$$M_{t,\lfloor mT \rfloor} = \lfloor \frac{m}{2} T \rfloor^{-1} \left(\sum_{i=1}^{\lfloor \frac{m}{2} T \rfloor} y_{t+i} - \sum_{i=1}^{\lfloor \frac{m}{2} T \rfloor} y_{t-i+1} \right) \tag{4.3}$$

which is the difference between the mean of the $\lfloor \frac{m}{2} T \rfloor$ observations $y_{t+1}, y_{t+2}, \ldots, y_{t+\lfloor \frac{m}{2} T \rfloor}$ and the mean of the $\lfloor \frac{m}{2} T \rfloor$ observations $y_t, y_{t-1}, \ldots, y_{t-\lfloor \frac{m}{2} T \rfloor + 1}$.

To ensure that at most only one level break can occur in the data spanned by any $M_{t,\lfloor mT \rfloor}$ we need to impose the restriction that $\tau_i - \tau_{i-1} \geq m$ for all $i = 2, \ldots, n$. Under this constraint, the DGP admits n level breaks occurring at unknown points across the interval Λ_T, with a minimum of $\lfloor mT \rfloor$ observations between breaks. As a result, n and m are bounded by the relation:

$$n \leq 1 + \lfloor \frac{\tau_U - \tau_L}{m} \rfloor =: n_{max}$$

which provides an upper bound for the maximum number of breaks assumed to be present for given choices of the *window width*, m, and the trimming parameters, τ_L and τ_U.

As HLT point out, $M_{\lfloor 0.5T \rfloor, \lfloor mT \rfloor}$ is the test suggested by maximum likelihood considerations in a stylized example with $m = 2/T$ in the $I(1)$ version of the example and with $m = 1$ in the $I(0)$ version. Hence they propose using a statistic that is based on $M_{t,\lfloor mT \rfloor}$ of (4.3) with $0 < m < 1$ for detecting possible multiple breaks in level even when the order of integration is unknown. Trying to detect the maximum function of $\left|M_{t,\lfloor mT \rfloor}\right|$ (notice that the signs of the possible breaks are also unknown) over all $t \in \Lambda_T$, i.e.

$$\mathcal{M} := \max_{t \in \Lambda_T} |M_{t,\lfloor mT \rfloor}|$$

For a given value of m, this statistic therefore takes the largest (in absolute value) fluctuation measure $|M_{t,\lfloor mT \rfloor}|$ over all possible points in Λ_T. Note that it is required that $\tau_L \geq m/2$ and $\tau_U \leq 1 - (m/2)$, so that $M_{t,\lfloor mT \rfloor}$ is only calculated from observed data.

The test statistic $\mathcal{M}$ is not pivotal, even under constant volatility, unless scaled by the long-run variance. In order to obtain a pivotal statistic in the case of constant volatility HLT derive estimates of the long-run variance. We now describe their procedure for how to estimate the long-run variance under $I(1)$ and $I(0)$ and present a robust procedure for testing for level breaks under both specifications. In section 4.3.1, we describe the procedure for estimating the long run variances and their behaviour under both $I(1)$ and $I(0)$ errors. In section 4.3.2, we describe the operational test against level breaks in model (4.1)-(4.2)) for the situation where the order of integration is unknown.

4.3.1 Estimation of the Long Run Variance of ε_t and u_t

We now describe the estimation of the long run variances ω_ε^2 that is relevant under $I(1)$ errors and ω_u^2 that is relevant under $I(0)$ errors. Initially, assume we know the order of integration.

The Long Run Variance of ε_t

First we consider estimating ω_ε^2 when the errors are known to be $I(1)$. It is obviously desirable from a power standpoint that the long run variance estimator is not influenced by the presence of the level breaks, bearing in mind that the number and timings of these breaks are unknown. Our first consideration is therefore estimation of the timing of the potential breaks. In the context of our reference level break model (4.1)-(4.2), we further assume that when there are n level breaks, that $|\gamma_1| > |\gamma_2| > \cdots > |\gamma_n|$. This ordering is adopted to expedite the arguments made below, and does not compromise the generality of the results.

Under Assumption $I(1)$, if the errors ε_t in (4.2) are Gaussian white noise and only one break is present (at time $\lfloor \tau_1 T \rfloor$), the optimal test of $\gamma^*_{1,T} = 0$ is based on the ML estimator $\Delta y_{\lfloor \tau_1 T \rfloor + 1}$. It makes sense, therefore, under $I(1)$ errors, to consider $|\Delta y_t|$ to identify any break points. Consequently, let $\hat{t}_1 := (\arg\max_{t \in \Omega_T} |\Delta y_t|) - 1$ where $\Omega_T := [\lfloor \tau_L T \rfloor + 1, \lfloor \tau_U T \rfloor + 1]$ (bearing in mind that the outliers are observed one observation after a corresponding break point). Next, since we are assuming that the breaks are separated by at least $\lfloor mT \rfloor$ observations, we now wish to exclude the dates $[\hat{t}_1 - \lfloor mT \rfloor + 1, \hat{t}_1 + \lfloor mT \rfloor - 1]$, so now let $\hat{t}_2 := \arg\max_{t \in \Omega_T - \Omega_{1,T}} |\Delta y_t|) - 1$ where $\Omega_{1,T} := [\hat{t}_1 - \lfloor mT \rfloor + 2, \hat{t}_1 + \lfloor mT \rfloor]$, then $\hat{t}_3 := (\arg\max_{t \in \Omega_T - \Omega_{1,T} - \Omega_{2,T}} |\Delta y_t|) - 1$ where $\Omega_{2,T} := [\hat{t}_2 - \lfloor mT \rfloor + 2, \hat{t}_2 + \lfloor mT \rfloor]$, and so on, until $\Omega_{\bar{n}+1,T} = \emptyset$. This procedure identifies $\bar{n}$ breaks points, where it can be shown that

$$\left\lfloor \frac{\lfloor \tau_U T \rfloor - \lfloor \tau_L T \rfloor + \lfloor mT \rfloor}{2\lfloor mT \rfloor - 1} \right\rfloor \leq \bar{n} \leq n_{max}. \tag{4.4}$$

Using the estimated break points, $\hat{t}_1, \ldots \hat{t}_{\bar{n}}$, we then remove the effect of the level breaks on the Δy_t series by taking the residuals $\hat{\varepsilon}_t$ from the OLS regression

$$\Delta y_t = \sum_{i=1}^{\bar{n}} \hat{\gamma}^*_i D_t(\hat{t}_i) + \hat{\varepsilon}_t, t = 2, \ldots, T \tag{4.5}$$

where the $D_t(\hat{t}_i) := \mathbb{I}(t = \hat{t}_i + 1), i = 1, \ldots, \bar{n}$, are one-time dummy variables. The Berk (1974)-type autoregressive spectral density estimator of ω^2_ε is the n obtained as

$$\bar{\omega}_\varepsilon^2 := \frac{\hat{\sigma}^2}{\hat{\pi}^2}$$

which is based on estimating the OLS regression

$$\Delta\hat{\varepsilon}_t = \hat{\pi}\hat{\varepsilon}_{t-1} + \sum_{j=1}^{k-1} \hat{\psi}_j \Delta\hat{\varepsilon}_{t-j} + \hat{e}_t, t = k+2, \ldots, T \tag{4.6}$$

with $\hat{\sigma}^2 := (T - 2k - 1)^{-1} \sum_{t=k+2}^{T} \hat{e}_t^2$. As is standard, we require that the lag truncation parameter, k, in (4.6) satisfies the condition that, as $T \to \infty$, $1/k + k^3/T \to 0$.

The Long Run Variance of u_t

Now consider estimating ω_u^2 in the case where the errors are known to be $I(0)$. Given the estimated break points, $\hat{t}_1, \ldots, \hat{t}_{\bar{n}}$, from the section above, we again account for the level breaks by taking the residuals $\hat{u}_t$ from the OLS regression

$$y_t = \hat{\alpha} + \sum_{i=1}^{\bar{n}} \hat{\gamma}_i^* DU_t(\hat{t}_i) + \hat{u}_t, t = 1, \ldots, T \tag{4.7}$$

where $DU_t(\hat{t}_i) := \mathbb{I}(t > \hat{t}_i)$, $i = 1, \ldots, \bar{n}$. The estimator of ω_u^2 in this case is given by

$$\hat{\omega}_u^2 := \frac{\hat{\sigma}^2}{\hat{\pi}^2}$$

where $\hat{\pi}$ and $\hat{\sigma}$ are now obtained from the OLS regression

$$\Delta\hat{u}_t = \hat{\pi}\hat{u}_{t-1} + \sum_{j=1}^{k-1}\hat{\psi}_j\Delta\hat{u}_{t-j} + \sum_{j=0}^{k-1}\sum_{i=1}^{\bar{n}}\hat{\psi}_{j,i}D_{t-j}(\hat{t}_i) + \hat{e}_t, t = k+1,\ldots,T, \tag{4.8}$$

with $\hat{\sigma}^2 := (T-(2+\bar{n})k)^{-1}\sum_{t=k+1}^{T}\hat{e}_t^2$, and where k again satisfies the condition that, as $T \to \infty$, $1/k + k^2/T \to 0$. Notice that, for the reasons outlined in Perron and Vogelsang (1992), the regression in (4.8) augments the usual ADF-type regression with the $\bar{n}$ one-time dummy variables, $D_t(\hat{t}_i)$, $i = 1,\ldots,\bar{n}$, and the $(k-1)$ lagged values of each of these.

4.3.2 The HLT tests

Having proposed suitable long run variance estimators we can finally define the feasible statistics proposed by HLT for detecting multiple level breaks. The asymptotic behavior of $\mathcal{M}$, along with the properties of the long run variance estimators, suggests the following statistics, appropriate under $I(1)$ and $I(0)$ errors, respectively:

$$S_1 := \hat{\omega}_{\varepsilon}^{-1}T^{-1/2}\mathcal{M} \tag{4.9}$$

$$S_0 := \hat{\omega}_{u}^{-1}T^{1/2}\mathcal{M} \tag{4.10}$$

Remark. It is useful for analysis in subsequent sections to note that S_1 and S_0 could equivalently be expressed as $S_1 := \max_{t\in\Lambda T} S_{1,t,\lfloor mT\rfloor}$ and $S_0 := \max_{t\in\Lambda T} S_{0,t,\lfloor mT\rfloor}$, where $S_{1,t,\lfloor mT\rfloor} := \hat{\omega}_{\varepsilon}^{-1}T^{-1/2}\left|M_{t,\lfloor mT\rfloor}\right|$ and

$$S_{0,t,\lfloor mT \rfloor} := \hat{\omega}_u^{-1} T^{-1/2} \left| M_{t,\lfloor mT \rfloor} \right|.$$

In the following theorem we establish the large sample behaviour of the S_1 and S_0 statistics of (4.9) and (4.10), respectively, in both $I(1)$ and $I(0)$ environments.

Theorem 2. *Let y_t be generated according to equations (4.1) and (4.2) and suppose Assumption $\mathcal{A}$ holds with $\omega(s) = \sigma$. Then,*

(a) Under Assumption I(1),

(i) $S_1 \xrightarrow{\omega} \sup_{r \in \Lambda} |L_1(r, m, c) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$;

(ii) $S_0 \xrightarrow{\omega} \frac{\sup_{r \in \Lambda} |L_1(r,m,c) + K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma})|}{Q^{1/2}(c,d,\tilde{\boldsymbol{\tau}})}$

(b) Under Assumption I(0),

(i) $S_1 = O_p(kT^{-1})$;

(ii) $S_0 \xrightarrow{\omega} \sup_{r \in \Lambda} |L_0(r, m) + K(r, m, \boldsymbol{\tau}, \boldsymbol{\gamma})|$

where $\boldsymbol{\tau} := [\tau_1, \tau_2, \ldots, \tau_n]$, $\boldsymbol{\gamma} := [\gamma_1, \gamma_2, \ldots, \gamma_n]$, $W(\cdot)$ *is a standard Brownian* *motion process on* $[0,1]$*, and*

$$L_0(r,m) := 2m^{-1}\left\{W(r+m/2) - 2W(r) + W(r-m/2)\right\}$$
$$L_1(r,m,c) := 2m^{-1}\left\{\int_r^{r+m/2} W_c(s)ds - \int_{r-m/2}^{r} W_c(s)ds\right\}$$
$$W_c(r) := \int_0^r e^{-(r-s)c} dW(s)$$
$$K(r,m,\boldsymbol{\tau},\boldsymbol{\gamma}) := \left\{\begin{array}{cc} 0 & \tau_L \le r \le \tau_1 - m/2 \\ \gamma_1\left(1 - \frac{|\tau-\tau_1|}{m/2}\right) & \tau_1 - m/2 \le r \le \tau_1 + m/2 \\ 0 & \tau_1 - m/2 \le r \le \tau_2 - m/2 \\ \gamma_2\left(1 - \frac{|\tau-\tau_2}{m/2}\right) & \tau_2 - m/2 \le r \le \tau_2 - m/2 \\ \vdots & \vdots \\ 0 & \tau_{n-1} + m/2 \le r \le \tau_n - m/2 \\ \gamma_n\left(1 - \frac{|\tau-\tau_n|}{m/2}\right) & \tau_n - m/2 \le r \le \tau_n + m/2 \\ 0 & \tau_n + m/2 \le r \le \tau_U \end{array}\right\}$$
$$Q(c,m,\tilde{\boldsymbol{\tau}}) := \frac{\left\{\int_0^1 H(r,c,m,\tilde{\boldsymbol{\tau}})^2 dr\right\}^2}{\left\{\int_0^1 H(r,c,m,\tilde{\boldsymbol{\tau}}) dW_c(r)\right\}^2}$$

where $H(r,c,m,\tilde{\tau})$ *is the (continuous-time) residual from the projection of the OU process,* $W_c(r)$*, onto the span of* $\{1, 1(r > \tilde{\tau}_1), \ldots, 1(r > \tilde{\tau}_{\tilde{n}})\}$ *and where* $\tilde{\tau}_i = \lim_{T\to\infty} T^{-1}\hat{t}_i$.

In order to test for level breaks in the absence of knowledge about the order of integration HLT develop a union of rejections procedure. Let cv_α^1 be the level α critical value for the distribution, $\sup_{r\in\Lambda} |L_1(r,m,c)|$, and let cv_α^0 be the level α critical value for the distribution, $\sup_{r\in\Lambda} |L_0(r,m)|$. These are the appropriate asymptotic critical values to use if the data is known to be I(1) or I(0), respectively. They propose the decision rule:

$$U: \text{ Reject } H_0 \text{ if either } S_1 > \kappa_\alpha cv_\alpha^1 \text{ or } S_0 > \kappa_\alpha cv_\alpha^0$$

Note that if $\kappa_\alpha = 1$ the procedure is correctly sized asymptotically under $I(0)$ errors since $S_1 \xrightarrow{p} 0$ in this case. But under $I(1)$ errors, $S_0 \stackrel{p}{\nrightarrow} 0$ and hence the procedure would be (slightly, as seen through simulations) over-sized with $\kappa_\alpha = 1$. Hence κ_α represents the adjustment needed in order to ensure the procedure is always correctly sized. See HLT or Harvey et al. (2009b) for more discussion.

4.4 Accounting for non-stationary volatility

The asymptotic distribution derived in the Theorem above is useful because under the null hypothesis it only depends on the (user-specified) window size, m, if the process is either stationary or exhibits a unit root. And hence the union of rejections procedure described above can be easily implemented without knowledge of any parameters of the model. However, the Theorem relies on the assumption that $\omega(s)$ is constant. If this is not the case, as we suspect in many applications, the asymptotic behavior described in the Theorem is not correct. Indeed, if the volatility is non-stationary then the asymptotic distribution of the test statistics may depend on the exact nature of this volatility process. For example, Cavaliere et al. (2011) show that the asymptotic distribution of a class of unit root tests depends on the following *variance profile* term:

$$\eta(s) = \frac{\int_0^s \omega(s)^2 ds}{\int_0^1 \omega(s)^2 ds}$$

In particular, the asymptotic distribution is distorted by $\eta(s)$. It is generated by the stochastic process $W(\eta(s))$ where $W(s)$ is a standard Brownian motion process on $[0, 1]$. A similar result holds for the statistics considered here, though this derivation is beyond the scope of this paper.

We performed a series of Monte Carlo exercises to explore the size distortion of the HLT test when the volatility of the driving process is not constant. We generated data from the model of equations (4.1) and (4.2) with $\alpha = \gamma_i = 0$, $\varepsilon_t = \eta_t$ and

$$\omega(s) = \begin{cases} \sigma_0 & \text{if } s \leq \tau_\sigma \\ \sigma_1 & \text{otherwise} \end{cases}$$

That is, we allowed the volatility to jump from σ_0 to σ_1 at time period $\lfloor \tau_\sigma T \rfloor$. We varied the ratio of σ_1/σ_0 across the four values $1/2, 1, 3, 5$. We performed the exercise for $\rho = 0$ and $\rho = 1$, i.e., an i.i.d. driving process and a pure random walk driving process. Tables 1A-1C demonstrate the performance of the HLT union of rejections procedure for this simulated data. The probability of (falsely) rejecting the null across 1000 Monte Carlo replications of the first set of exercises is recorded in the rows labeled $\rho = 0$. The results of the second set of exercises are labeled $\rho = 1$. We repeated the exercises for sample sizes $T = 150, 300, 600$, and 1200. In all of the simulations we performed the tests using the finite sample critical values reported in Harvey et al. (2010), and not the asymptotic critical values, as they find that there are noticeable differences, particularly for smaller sample sizes. Table 1A reports the results when $\tau_\sigma = .25$ and Tables 1B and 1C repeat these simulations but now varying the period in which the break occurs to be either after half or three-quarters of the periods pass.

In general the failure to properly account for the break in volatility leads to over-rejection of the null when a level break is not present. This phenomenon is worse when the jump is larger and when it occurs closer to the middle of the time period observed. The rejection rate is also typically worse when the underlying process is stationary, all else held constant. Finally, the tests with larger values of m are evidently less affected. Recall that m refers to the size of the window used to detect fluctuations in the data. This is expected because if the window is larger it is easier to see whether a sudden jump should be attributed to a level break or a jump in the volatility of the process.

As in Cavaliere and Taylor (2008) and Cavaliere et al. (2011) we implement a wild bootstrap procedure to account for this problem. As we will discuss later it is necessary to develop two different wild bootstrap-based resampling schemes, one for stationary processes and one for non-stationary processes.

4.4.1 Wild bootstrap procedure I

We first describe a straightforward procedure that generates bootstrap samples that can be used to approximate the distribution of S_1 and S_0 under the null hypothesis of no level breaks if the process is $I(1)$. If the process is known to be $I(1)$ then an appropriate testing procedure would be to reject if S_1 exceeds the $1-\alpha$ percentile of the distribution, as calculated using many draws from this bootstrap distribution. The first step is to remove the level breaks from the series.

(a) Estimate the break points $\hat{t}_i$ for $i = 1, \ldots, \bar{n}$ as described in Section 4.3.

(b) Estimate the regression

$$y_t - y_{t-1} = \sum_{i=1}^{\bar{n}} \hat{\gamma}_i D_t(\hat{t}_i) + \hat{\varepsilon}_t$$

via OLS where $D_t(s) = \mathbf{1}(t = s + 1)$.

(c) Remove the break:

$$\tilde{u}_t = y_t - \sum_{i=1}^{\bar{n}} \hat{\gamma}_i DU_t(\hat{t}_i)$$

where $DU_t(s) = \mathbf{1}(t > s)$.

The next step is to estimate the AR(1) coefficient. To do this, regress $\tilde{u}_t$ on $\tilde{u}_{t-1}$. That is, estimate the regression

$$\tilde{u}_t = \hat{\alpha} + \hat{\rho}\tilde{u}_{t-1} + \hat{\varepsilon}_t$$

Now we can use the residuals from this regression to construct a wild bootstrap sample. Let $w_1, \ldots, w_T$ be i.i.d. $\sim \mathcal{N}(0,1)$. Then define

$$\varepsilon_t^* = w_t \hat{\varepsilon}_t$$

Then we construct a new sample of y_t's that imposes the null of no level break.

$$y_1^* = \varepsilon_1^*$$
$$y_t^* = \hat{\rho} y_{t-1}^* + \varepsilon_t^*, t = 2, \dots, T$$

In order to simulate the wild bootstrap critical values for a test statistic we have to create such a sample N_{boot} times. Each time we record the value of the test statistic computed on the bootstrap sample, S_j^*. Then the bootstrap critical value is the $1-\alpha$ quantile of the sample $S_1^*, \dots, S_{N_{boot}}^*$.

If the process is stationary then this procedure will not be accurate for reasons explained below. We now describe a procedure that generates bootstrap samples that can be used to approximate the distribution of S_1 and S_0 under the null hypothesis of no level breaks if the process is $I(0)$. Of course if the process is known to be $I(0)$ then an appropriate testing procedure would be to reject if S_0 exceeds the $1-\alpha$ percentile of the distribution, as calculated using many draws from this second bootstrap distribution.

4.4.2 Wild bootstrap procedure II

(a) Estimate the break points $\hat{t}_i$ for $i = 1, \dots, \bar{n}$ as described in Section 4.3.

(b) Estimate the regression

$$y_t = \hat{\alpha} + \hat{\rho} y_{t-1} + \sum_{i=1}^{\bar{n}} \hat{\gamma}_i D_t(\hat{t}_i) + \sum_{i=1}^{\bar{n}} \hat{\gamma}_i DU_t(\hat{t}_i) + \hat{\varepsilon}_t$$

and generate the residuals, $\hat{\varepsilon}_t$.

(c) Let $w_1, \ldots, w_T$ be i.i.d. $\sim \mathcal{N}(0,1)$. Then define

$$\varepsilon_t^* = w_t \hat{\varepsilon}_t$$

(d) Construct a new sample of y_t's that imposes the null of no level break.

$$y_1^* = \varepsilon_1^*$$
$$y_t^* = \hat{\alpha} + \hat{\rho} y_{t-1}^* + \varepsilon_t^*, t = 2, \ldots, T$$

4.4.3 Difficulty of bootstrap for the U test

Now note that if the order of integration is unknown and the volatility is possibly non-constant we neither know which test statistic to use – S_1 or S_0 – nor what boostrap procedure to use to derive critical values. As shown in HLT for the constant volatility case, the S_1 and S_0 tests will be under-sized under the $I(0)$ and $I(1)$ assumptions, respectively, i.e. under the "wrong" assumptions. This remains the case for the more general model here. So we would prefer to implement a union of rejections type test, as HLT do. The HLT statistic $U = \max\{S_1, \frac{cv_1}{cv_0} S_0\}$ makes sense because

$$U \leq cv_1 \Leftrightarrow S_0 \leq cv_0$$

asymptotically, under $I(0)$, since in this case $S_1 \xrightarrow{p} 0$. Then we can adjust the critical value from cv_1 to κcv_1 to control the size in the $I(1)$ case.

In our more general model the critical values cv_1 and cv_0 are not known and possibly depend on unknown parameters of the volatility process. Hence if we instead use the bootstrap critical values of S_1 and S_0 to construct U, the HLT logic no longer holds. Another possibility is to use the asymptotic critical values (under the assumption of constant volatility, from HLT), which is what we do.

4.4.4 The wild bootstrapped U test for level breaks

Following the discussion in the previous paragraph we define the test statistic

$$U(w) = max\{S_1, wS_0\},$$

which can be computed from data $y_1, \ldots, y_T$ for any scalar w where S_1 and S_0 are defined above. Here w describes how S_0 is weighted relative to S_1.

The proposed test rejects the null of no level breaks if $U(w) > \hat{c}^*_\alpha(w)$ where

$$\hat{c}^*_\alpha(w) = \max\{\hat{c}^*_{1,\alpha}(w), \hat{c}^*_{2,\alpha}(w)\}$$

and the critical values $\hat{c}^*_{j,\alpha}(w)$ for $j = 1, 2$ are obtained via the wild bootstrap as follows.

(a) Set the number of boostrap samples to N and fix w.

(b) Obtain a sample $y^*_1, \ldots, y^*_T$ using Wild bootstrap procedure I if $j = 1$ and Wild bootstrap procedure II if $j = 2$. Calculate $U(w)$ using this new sample and call it $U^*_1(w)$.

(c) Repeat the previous step for $i = 2, \dots, N$ obtaining $U_2^*(w), \dots, U_N^*(w)$.

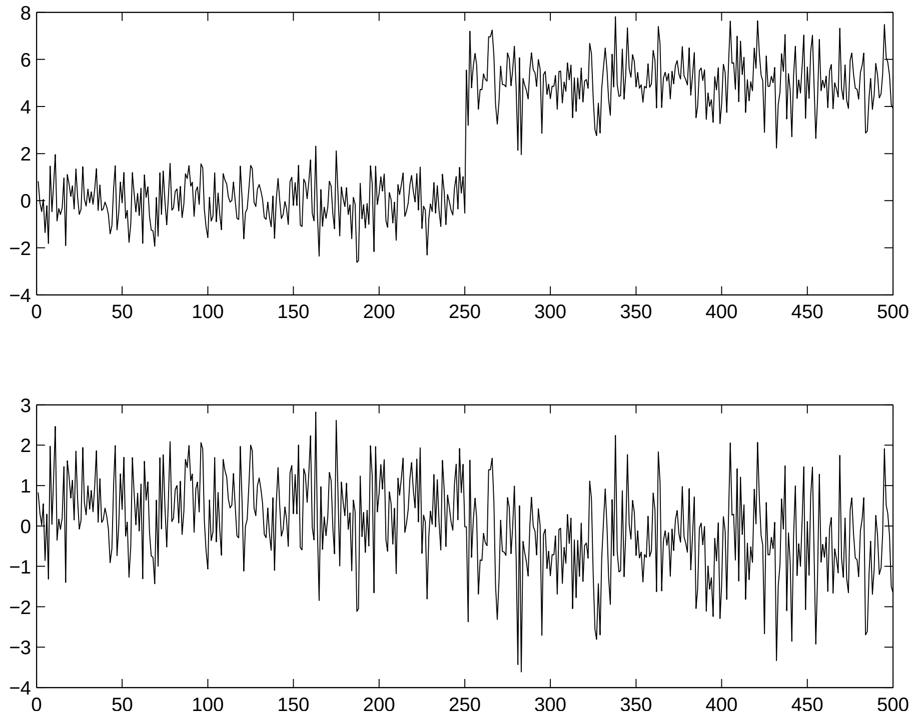
(d) Define $\hat{c}_{j,\alpha}^*(w) = \inf\{t : N^{-1} \sum_{i=1}^N \mathbf{1}(U_i^*(w) \geq t) \leq \alpha\}$.

This procedure can be performed for any value of the weighting parameter w. Ideally w should be equal to the critical value of S_1 under the I(1) assumption divided by the critical value of S_0 under the I(0) assumption. These critical values cannot both be accurately approximated using the bootstrap procedures. Instead we use the asymptotic critical values derived under the assumption of constant volatility, taken from HLT.

4.4.5 Problem with wild bootstrap procedure I

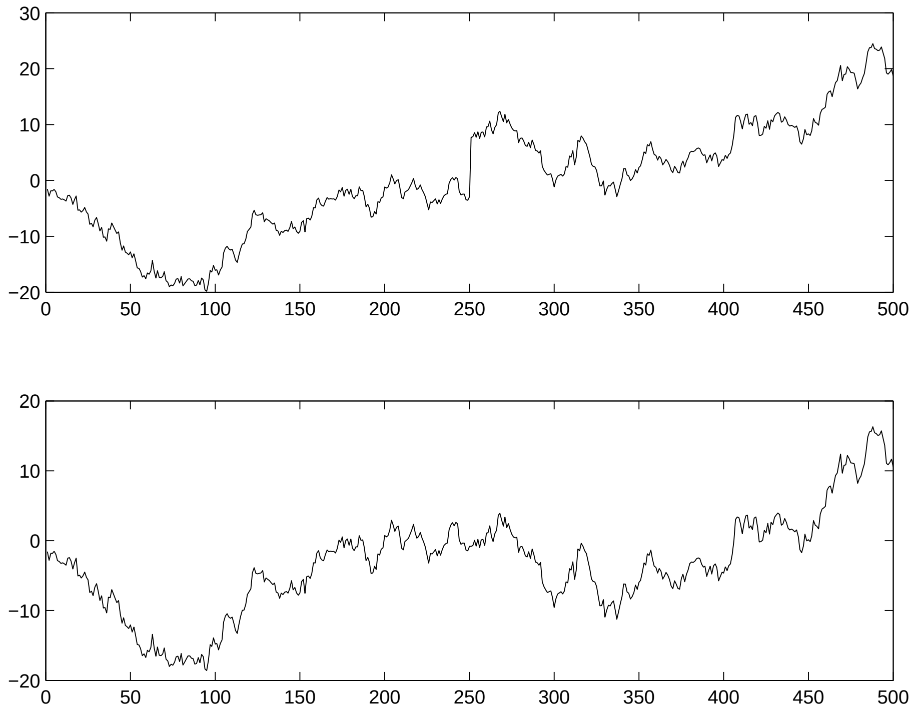
It may not be obvious why the resampling scheme given by wild bootstrap procedure I does not also work when the underlying process is stationary. The logic behind the test is as follows. If the data exhibit a level break then we want to estimate the break and remove it from the data. Then we use the wild bootstrap on the adjusted data to approximate the distribution of the test statistic U under the null hypothesis but with the correct volatility process. If the data is generated by a model with no break in level we would want the procedure to still approximate the distribution of U under the null hypothesis of no break in level. It turns out that this works fairly well for a $I(1)$ process but not for a stationary process. First, look at what it does for a process with a level break. We generated an i.i.d. $N(0, 1)$ process with a break in level at observation 250 out of 500. Figure 1(a) below shows the original process and Figure 1(b) shows the process obtained by using wild bootstrap procedure I, except that we take $\varepsilon_t^* = \hat{\varepsilon}_t$.

Figure 1. (a) I(0) process with level break, (b) bootstrapped residuals, procedure I



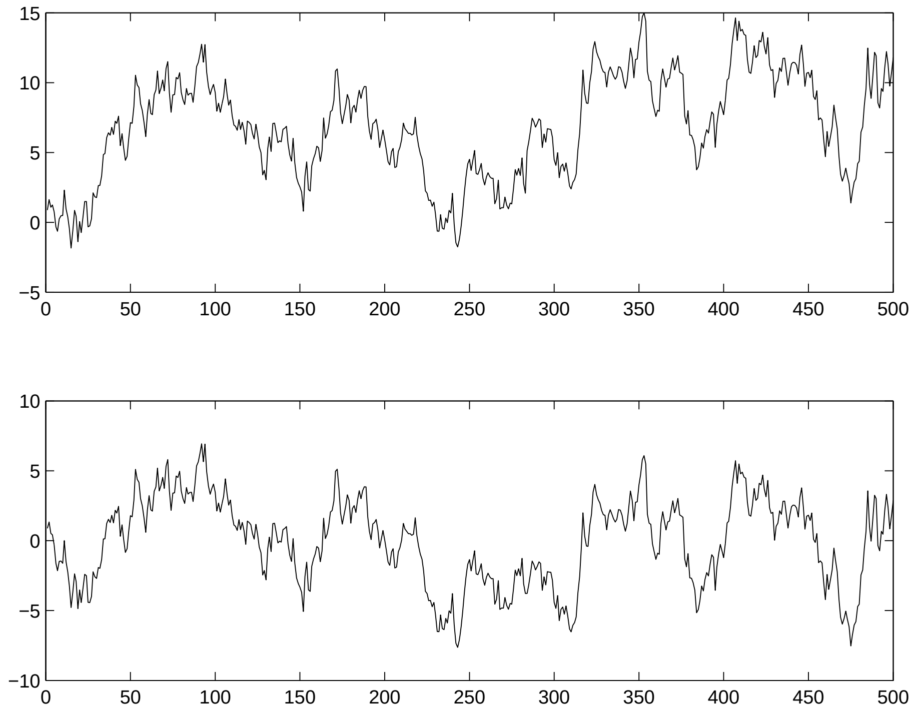
This is what we would like to see. The process removes the level break and does not seem to otherwise distort the time series behavior of the series. Next, consider what happens if the process is non-stationary with a break. We generate $u_t = u_{t-1} + \varepsilon_t$ where ε_t is i.i.d. $N(0, 1)$ and then add a break in level to u_t at observation 250 out of 500. The original series and the bootstrapped residuals are shown in Figure 2.

Figure 2. (a) I(1) process with level break, (b) bootstrapped residuals, procedure I



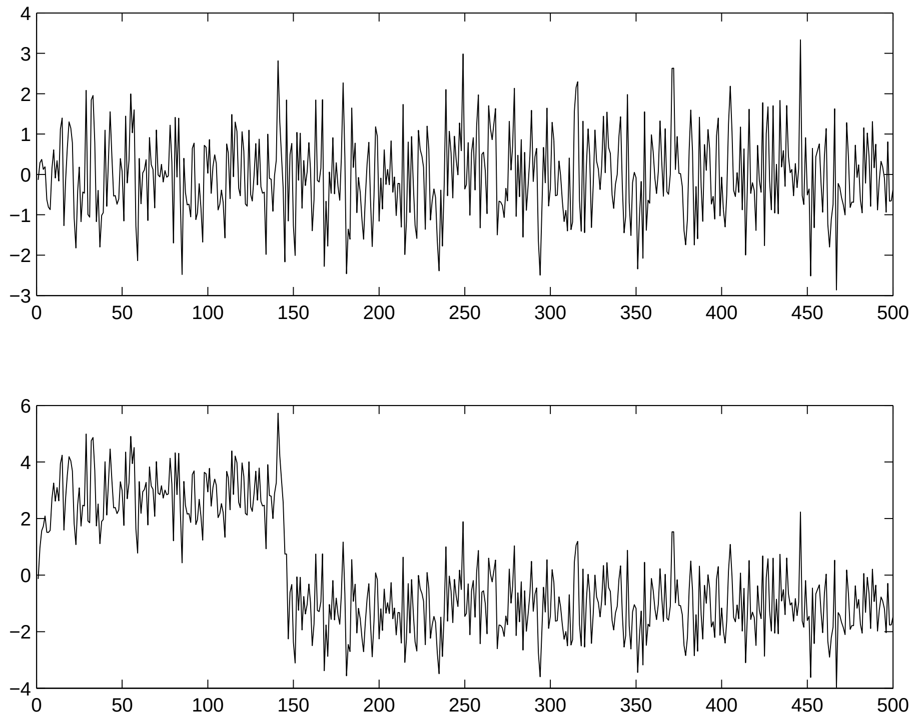
First note that the break is slightly less noticeable than in the previous figure because the process is non-stationary and the simulated break was of the same magnitude. But it still apparent that the procedure removed the break from the data properly. Next, consider what happens if the process is non-stationary with no break. Here we generated the data just as in the previous figure except that a break in level was not added. The resulting series are shown in Figure 3.

Figure 3. (a) I(1) process without level break, (b) bootstrapped residuals, procedure I



The procedure produces a time series that does not look distorted in any way. In particular, the new series does not appear to have a level break. Finally, consider what happens when the process is stationary with no break. Using an i.i.d. normal data generation process as in Figure 1, but this time without a level break, we plot the resulting series in Figure 4.

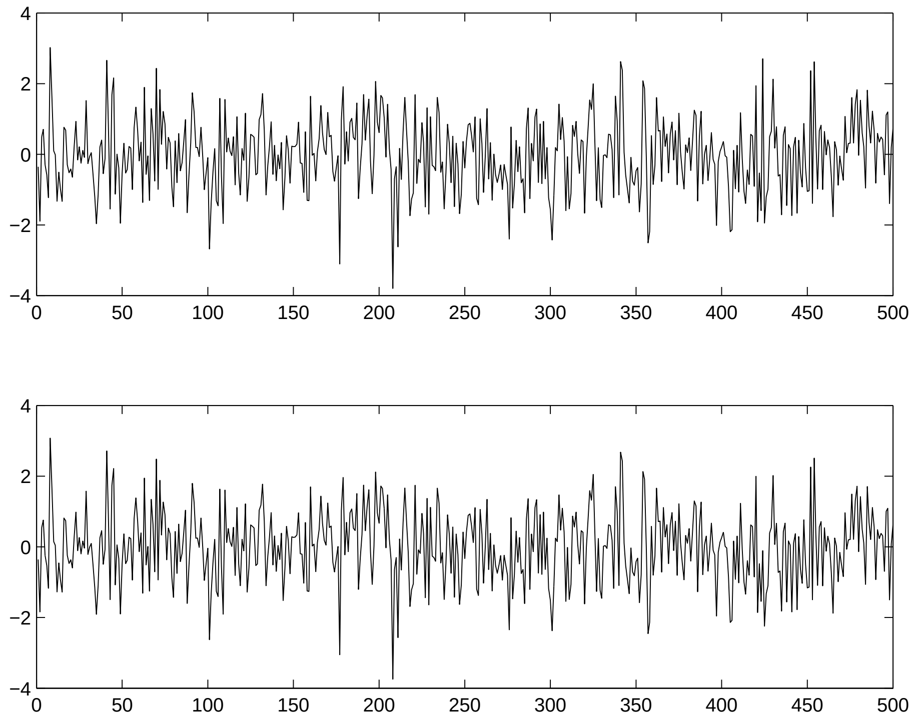
Figure 4. (a) I(0) process without level break,(b) bootstrapped residuals, procedure I



This example is where a problem occurs. The main problem is that the estimate of the break magnitude(s), $\hat{\gamma}_i$, are not close to zero when the data is generated by a model with no breaks. The differencing is what throws this off. For example, regressing on a single dummy $D_t(\hat{t})$ will result in $\hat{\gamma} = \max \Delta y_t$. On the other hand, regressing levels of y_t on $DU_t(\hat{t})$ gives $\hat{\gamma} = \bar{y}_{t \geq \hat{t}} - \bar{y}_{t \leq \hat{t}}$.

Now in Figure 5 we show what happens when the process is stationary with no break using the wild bootstrap II procedure.

Figure 5. (a) I(0) process without level break, (b) bootstrapped residuals, procedure II



As expected, the resampling no longer creates a level break in the series, nor does it distort the process in any other noticeable way. Hence wild bootstrap procedure II is valid for stationary data, whereas wild bootstrap procedure I is not.

Because bootstrap procedure I exhibits this problem when the data is I(0) we also rely on the alternative bootstrap procedure II. By using the larger of the two resulting critical values we ensure that the testing procedure will not be oversized. However, depending on the direction of failure in procedure I for I(0) data the procedure could potentially be very conservative, or undersized. We find in the simulations in the subsequent section that this is generally not the case.

4.4.6 Simulation results

We now present results of simulations that show the size and power of the bootstrap-based testing procedure under different data-generating processes.

In Tables 2A-2C we assess the size of the bootstrap-based testing procedure. The number of Monte Carlo simulations is now reduced from 1000 to 500 as they are computationally expensive. The same set of data-generating processes are used except that we now included the intermediate $\rho = 0.95$ case and only use samples of size $T = 150$ or 300. Across the board, the size is much closer to the nominal size, 0.05, than the size of the HLT test. The worst performance is when m is large, $\rho = 1$ and the volatility break occurs toward the beginning or end of the time period observed. Even in this case though the size is lower (closer to the nominal size) than the size of the HLT test.

Finally, Tables 3A-3D list the power against several different alternatives for these same data-generating processes when $T = 150$. Tables 3A-3C report the power from exercises in which the data was generated with a break in level after half of the periods have passed of two different magnitudes, $\gamma = 5$ and $\gamma = 10$. The simulations recorded in Table 3D used three level breaks of equal size - $\gamma = 5$ in the first panel and $\gamma = 10$ in the second panel - occurring after 1/4, 1/2 and 3/4 of the time periods.

First we can compare the constant volatility rows in the $\gamma = 5$ panel of Table 3B with Table 8 in HLT. Note that our bootstrap-based test maintains high power under constant volatility when $\rho = 0$. Furthermore, the power of our test under constant volatility when the underlying process has a unit root or is near unit root is very similar to that of the HLT test, roughly ranging from .07 to .13 across different values of m. Comparing the $\gamma = 10$ panels to Table 7 in HLT we see that the power of our test converges more quickly toward 1.

It is also clear from these results that our test typically maintains similar levels of local power when there is a small jump in volatility, though for a fixed γ the power suffers for larger jumps. The timing of the break in volatility also seems to be an important determinant of the power of the test, though the pattern is hard to decipher. Finally, it is also evident that using a larger window size, m, for the test leads to smaller losses of power. This is an important discovery because when volatility is constant both our test and the HLT test generally have higher power when m is smaller. Thus there is an interesting tradeoff between (i) using a wider window to distinguish between level breaks and volatility jumps and (ii) using a smaller window to better distinguish level breaks from non-stationarity in the underlying process.

The results in Table 3D mostly support this conclusion that when a time series exhibits jumps in volatility our bootstrapped version of HLT's testing procedure can still detect level breaks with similar power to the HLT test when there is no jump in volatility. However, when both γ - the level break magnitude - and the size of the jump volatility are large this does not appear to be true. In the last panel of Table 3D we find that the power is quite low both for the stationary and non-stationary cases. Note however, that for a window size of $m = .30$ test is inconsistent as there is now more than one break within the window, violating the requirement of at most one break.

4.5 Conclusion

Harvey et al. (2010) discuss a class of test statistics for detecting multiple level breaks in a time series and propose a testing procedure that is valid regardless of the order of integration of the underlying process. In this paper we studied the behavior of their tests when the driving process does not exhibit stationary volatility. We first demonstrated through a series of Monte Carlo exercises that the union of rejections procedure proposed by HLT is oversized when there is a break in the volatility. The size distortion is worse when the jump is larger, when it occurs closer to the middle of the time period observed, and when the data generating process satisfies Assumption I(0). The size distortion is also worse when a smaller window is used for the test.

Following Cavaliere and Taylor (2008) and Cavaliere et al. (2011) we propose a procedure based on the wild boostrap to perform valid tests when the underlying process has non-stationary volatility. We use the HLT test statistics and employ a resampling procedure to obtain valid critical values. Since we consider the order of integration to be unknown we are forced to use two separate bootstrap procedures to obtain two separate critical values, each of which is valid under the corresponding stationarity assumption. We reject the null hypothesis if the statistic exceeds both critical values.

We find through a Monte Carlo study that this adjusted testing procedure controls the size fairly well, even in small samples. We also find that while the power of the testing procedure is similar to the power of the HLT procedure when the volatility is constant in some cases where the volatility jump is small, the power is considerably lower for larger volatility jumps. We leave the theoretical derivation of the asymptotic behavior of the tests under non-stochastic volatility to future research. In future research we also hope to consider adapting the test statistic itself to account for changes in the volatility, in the hope of improving the power of the testing procedure.

# of Monte Carlo Simulations: 1000

Table 1.A Finite sample sizes of nominal 0.05-level tests: normal innovations with volatility breaks

τσ=0.25

		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
Panel A. T=150						
U						
constant volatility	ρ = 1	0.049	0.041	0.042	0.041	0.039
	ρ = 0	0.025	0.018	0.009	0.013	0.006
σ1/σ0=0.5	ρ = 1	0.121	0.112	0.109	0.109	0.11
	ρ = 0	0.121	0.088	0.064	0.045	0.036
σ1/σ0=3	ρ = 1	0.121	0.099	0.081	0.077	0.068
	ρ = 0	0.081	0.063	0.046	0.029	0.027
σ1/σ0=5	ρ = 1	0.142	0.107	0.088	0.08	0.072
	ρ = 0	0.087	0.072	0.051	0.031	0.029
Panel B. T=300						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.051	0.04	0.045	0.04	0.045
	ρ = 0	0.051	0.031	0.027	0.012	0.012
σ1/σ0=0.5	ρ = 1	0.116	0.113	0.115	0.123	0.115
	ρ = 0	0.152	0.103	0.079	0.061	0.048
σ1/σ0=3	ρ = 1	0.133	0.09	0.092	0.09	0.07
	ρ = 0	0.134	0.107	0.075	0.051	0.035
σ1/σ0=5	ρ = 1	0.141	0.101	0.098	0.095	0.082
	ρ = 0	0.152	0.12	0.081	0.059	0.039
Panel C. T=600						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.044	0.044	0.05	0.051	0.059
	ρ = 0	0.046	0.037	0.029	0.017	0.016
σ1/σ0=0.5	ρ = 1	0.123	0.112	0.113	0.122	0.114
	ρ = 0	0.179	0.125	0.077	0.057	0.052
σ1/σ0=3	ρ = 1	0.141	0.127	0.108	0.093	0.094
	ρ = 0	0.147	0.121	0.099	0.058	0.04
σ1/σ0=5	ρ = 1	0.149	0.141	0.127	0.107	0.099
	ρ = 0	0.161	0.131	0.105	0.063	0.048
Panel D. T=1200						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.045	0.05	0.054	0.046	0.055
	ρ = 0	0.055	0.039	0.043	0.027	0.021
σ1/σ0=0.5	ρ = 1	0.126	0.113	0.11	0.117	0.113
	ρ = 0	0.188	0.128	0.092	0.06	0.048
σ1/σ0=3	ρ = 1	0.128	0.127	0.112	0.096	0.083
	ρ = 0	0.137	0.11	0.096	0.07	0.055
σ1/σ0=5	ρ = 1	0.137	0.136	0.12	0.106	0.089
	ρ = 0	0.147	0.12	0.108	0.078	0.06

# of Monte Carlo Simulations: 1000

Table 1.B Finite sample sizes of nominal 0.05-level tests: normal innovations with volatility breaks

τσ=0.5

		Panel A. T=150				
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.049	0.041	0.042	0.041	0.039
	ρ = 0	0.025	0.018	0.009	0.013	0.006
σ1/σ0=0.5	ρ = 1	0.165	0.149	0.123	0.113	0.099
	ρ = 0	0.133	0.105	0.075	0.059	0.045
σ1/σ0=3	ρ = 1	0.212	0.162	0.148	0.137	0.121
	ρ = 0	0.21	0.144	0.116	0.086	0.055
σ1/σ0=5	ρ = 1	0.269	0.204	0.187	0.178	0.153
	ρ = 0	0.251	0.18	0.145	0.104	0.076
		Panel B. T=300				
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.051	0.04	0.045	0.04	0.045
	ρ = 0	0.051	0.031	0.027	0.012	0.012
σ1/σ0=0.5	ρ = 1	0.171	0.151	0.137	0.135	0.114
	ρ = 0	0.173	0.152	0.102	0.067	0.052
σ1/σ0=3	ρ = 1	0.221	0.164	0.154	0.139	0.116
	ρ = 0	0.274	0.229	0.172	0.124	0.092
σ1/σ0=5	ρ = 1	0.286	0.22	0.205	0.187	0.152
	ρ = 0	0.325	0.273	0.205	0.149	0.114
		Panel C. T=600				
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.044	0.044	0.05	0.051	0.059
	ρ = 0	0.046	0.037	0.029	0.017	0.016
σ1/σ0=0.5	ρ = 1	0.157	0.127	0.127	0.124	0.108
	ρ = 0	0.214	0.168	0.121	0.086	0.069
σ1/σ0=3	ρ = 1	0.23	0.209	0.172	0.172	0.154
	ρ = 0	0.316	0.233	0.179	0.14	0.11
σ1/σ0=5	ρ = 1	0.277	0.257	0.219	0.213	0.188
	ρ = 0	0.375	0.287	0.218	0.17	0.129
		Panel D. T=1200				
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.045	0.05	0.054	0.046	0.055
	ρ = 0	0.055	0.039	0.043	0.027	0.021
σ1/σ0=0.5	ρ = 1	0.155	0.141	0.134	0.123	0.12
	ρ = 0	0.208	0.157	0.107	0.084	0.075
σ1/σ0=3	ρ = 1	0.216	0.208	0.185	0.164	0.14
	ρ = 0	0.294	0.239	0.19	0.139	0.112
σ1/σ0=5	ρ = 1	0.26	0.25	0.226	0.197	0.169
	ρ = 0	0.36	0.285	0.226	0.167	0.136

# of Monte Carlo Simulations: 1000

Table 1.C Finite sample sizes of nominal 0.05-level tests: normal innovations with volatility breaks

τσ=0.75

Panel A. T=150						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.049	0.041	0.042	0.041	0.039
	ρ = 0	0.025	0.018	0.009	0.013	0.006
σ1/σ0=0.5	ρ = 1	0.116	0.09	0.081	0.073	0.076
	ρ = 0	0.072	0.048	0.036	0.036	0.02
σ1/σ0=3	ρ = 1	0.221	0.173	0.183	0.185	0.191
	ρ = 0	0.223	0.154	0.104	0.091	0.078
σ1/σ0=5	ρ = 1	0.308	0.247	0.258	0.282	0.303
	ρ = 0	0.334	0.228	0.163	0.141	0.127
Panel B. T=300						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.051	0.04	0.045	0.04	0.045
	ρ = 0	0.051	0.031	0.027	0.012	0.012
σ1/σ0=0.5	ρ = 1	0.108	0.095	0.081	0.078	0.071
	ρ = 0	0.103	0.084	0.053	0.03	0.027
σ1/σ0=3	ρ = 1	0.217	0.199	0.19	0.205	0.203
	ρ = 0	0.277	0.228	0.163	0.115	0.08
σ1/σ0=5	ρ = 1	0.306	0.283	0.279	0.279	0.291
	ρ = 0	0.38	0.301	0.218	0.162	0.125
Panel C. T=600						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.044	0.044	0.05	0.051	0.059
	ρ = 0	0.046	0.037	0.029	0.017	0.016
σ1/σ0=0.5	ρ = 1	0.109	0.1	0.087	0.086	0.082
	ρ = 0	0.133	0.094	0.065	0.046	0.038
σ1/σ0=3	ρ = 1	0.204	0.171	0.167	0.18	0.183
	ρ = 0	0.299	0.229	0.174	0.123	0.088
σ1/σ0=5	ρ = 1	0.293	0.253	0.264	0.286	0.285
	ρ = 0	0.409	0.308	0.233	0.183	0.136
Panel D. T=1200						
U						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.045	0.05	0.054	0.046	0.055
	ρ = 0	0.055	0.039	0.043	0.027	0.021
σ1/σ0=0.5	ρ = 1	0.111	0.098	0.085	0.085	0.09
	ρ = 0	0.12	0.088	0.068	0.067	0.047
σ1/σ0=3	ρ = 1	0.23	0.203	0.191	0.184	0.19
	ρ = 0	0.34	0.253	0.185	0.139	0.124
σ1/σ0=5	ρ = 1	0.321	0.275	0.268	0.276	0.284
	ρ = 0	0.45	0.333	0.261	0.207	0.172

# of Monte Carlo Simulations: 500
# of Bootstraps: 100

Table 2.A Finite sample sizes of nominal 5%-level tests: normal innovations with volatility breaks, WBS

τσ=0.25

Panel A. T=150

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.042	0.064	0.072	0.084	0.088
	ρ = 0.95	0.034	0.054	0.048	0.056	0.046
	ρ = 0	0.03	0.046	0.044	0.052	0.038
σ1/σ0=3	ρ = 1	0.014	0.044	0.066	0.06	0.066
	ρ = 0.95	0.002	0.038	0.052	0.052	0.058
	ρ = 0	0.008	0.02	0.04	0.044	0.04
σ1/σ0=5	ρ = 1	0.012	0.04	0.072	0.066	0.064
	ρ = 0.95	0.004	0.04	0.054	0.052	0.06
	ρ = 0	0.006	0.018	0.04	0.044	0.032

Panel B. T=300

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.036	0.048	0.072	0.08	0.078
	ρ = 0.95	0.016	0.028	0.024	0.024	0.042
	ρ = 0	0.026	0.038	0.038	0.036	0.04
σ1/σ0=0.5	ρ = 1	0.044	0.062	0.08	0.09	0.096
	ρ = 0.95	0.026	0.018	0.036	0.046	0.044
	ρ = 0	0.05	0.058	0.056	0.048	0.05
σ1/σ0=3	ρ = 1	0.036	0.044	0.052	0.054	0.064
	ρ = 0.95	0.02	0.022	0.024	0.034	0.036
	ρ = 0	0.016	0.04	0.036	0.038	0.034
σ1/σ0=5	ρ = 1	0.04	0.044	0.052	0.058	0.062
	ρ = 0.95	0.018	0.024	0.024	0.034	0.044
	ρ = 0	0.016	0.04	0.03	0.034	0.034

# of Monte Carlo Simulations: 500
# of Bootstraps: 100

Table 2.B Finite sample sizes of nominal 5%-level tests: normal innovations with volatility breaks, WBS

τσ=0.5

Panel A. T=150

U		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.024	0.044	0.076	0.086	0.098
	ρ = 0.95	0.022	0.038	0.024	0.044	0.042
	ρ = 0	0.016	0.028	0.042	0.038	0.044
σ1/σ0=3	ρ = 1	0.018	0.052	0.066	0.074	0.078
	ρ = 0.95	0.014	0.044	0.056	0.066	0.062
	ρ = 0	0.008	0.02	0.056	0.05	0.042
σ1/σ0=5	ρ = 1	0.024	0.042	0.066	0.078	0.076
	ρ = 0.95	0.012	0.046	0.066	0.068	0.058
	ρ = 0	0.008	0.024	0.048	0.042	0.038

Panel B. T=300

U		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.036	0.048	0.072	0.08	0.078
	ρ = 0.95	0.016	0.028	0.024	0.024	0.042
	ρ = 0	0.026	0.038	0.038	0.036	0.04
σ1/σ0=0.5	ρ = 1	0.058	0.076	0.084	0.064	0.066
	ρ = 0.95	0.03	0.042	0.042	0.04	0.036
	ρ = 0	0.038	0.052	0.048	0.056	0.052
σ1/σ0=3	ρ = 1	0.03	0.048	0.06	0.048	0.042
	ρ = 0.95	0.022	0.028	0.03	0.036	0.05
	ρ = 0	0.03	0.036	0.06	0.044	0.044
σ1/σ0=5	ρ = 1	0.03	0.05	0.058	0.052	0.058
	ρ = 0.95	0.024	0.03	0.03	0.044	0.052
	ρ = 0	0.026	0.03	0.05	0.046	0.048

# of Monte Carlo Simulations: 500
# of Bootstraps: 100

Table 2.C Finite sample sizes of nominal 5%-level tests: normal innovations with volatility breaks, WBS

τσ=0.75

Panel A. T=150

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.018	0.038	0.06	0.076	0.092
	ρ = 0.95	0.01	0.028	0.036	0.048	0.058
	ρ = 0	0.004	0.018	0.03	0.022	0.03
σ1/σ0=3	ρ = 1	0.052	0.086	0.096	0.094	0.082
	ρ = 0.95	0.06	0.068	0.078	0.082	0.1
	ρ = 0	0.034	0.068	0.068	0.07	0.062
σ1/σ0=5	ρ = 1	0.068	0.088	0.094	0.098	0.098
	ρ = 0.95	0.074	0.076	0.078	0.086	0.096
	ρ = 0	0.034	0.064	0.06	0.054	0.034

Panel B. T=300

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.036	0.048	0.072	0.08	0.078
	ρ = 0.95	0.016	0.028	0.024	0.024	0.042
	ρ = 0	0.026	0.038	0.038	0.036	0.04
σ1/σ0=0.5	ρ = 1	0.05	0.062	0.078	0.076	0.094
	ρ = 0.95	0.014	0.03	0.018	0.018	0.038
	ρ = 0	0.02	0.04	0.04	0.054	0.054
σ1/σ0=3	ρ = 1	0.064	0.078	0.074	0.07	0.076
	ρ = 0.95	0.028	0.046	0.054	0.046	0.056
	ρ = 0	0.06	0.074	0.052	0.034	0.032
σ1/σ0=5	ρ = 1	0.062	0.076	0.08	0.082	0.078
	ρ = 0.95	0.038	0.05	0.056	0.062	0.062
	ρ = 0	0.058	0.064	0.052	0.032	0.028

# of Monte Carlo Simulations: 500
# of Bootstraps: 100
One vol break at: τσ=0.25
One level break at: τ=0.5

Table 3.A Finite sample powers of nominal 0.05-level tests: normal innovations with a single volatility break, WBS

Panel A. T=150, γ= 0 (No level break)						
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.042	0.064	0.072	0.084	0.088
	ρ = 0.95	0.034	0.054	0.048	0.056	0.046
	ρ = 0	0.03	0.046	0.044	0.052	0.038
σ1/σ0=3	ρ = 1	0.014	0.044	0.066	0.06	0.066
	ρ = 0.95	0.002	0.038	0.052	0.052	0.058
	ρ = 0	0.008	0.02	0.04	0.044	0.04
σ1/σ0=5	ρ = 1	0.012	0.04	0.072	0.066	0.064
	ρ = 0.95	0.004	0.04	0.054	0.052	0.06
	ρ = 0	0.006	0.018	0.04	0.044	0.032
Panel B. T=150, γ= 5 (one level break)						
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.08	0.11	0.096	0.112	0.126
	ρ = 0.95	0.11	0.114	0.13	0.19	0.232
	ρ = 0	0.966	0.984	0.964	0.946	0.9
σ1/σ0=0.5	ρ = 1	0.386	0.296	0.242	0.248	0.274
	ρ = 0.95	0.464	0.39	0.32	0.36	0.438
	ρ = 0	1	1	1	1	0.992
σ1/σ0=3	ρ = 1	0.014	0.042	0.066	0.068	0.074
	ρ = 0.95	0.012	0.032	0.044	0.066	0.058
	ρ = 0	0.174	0.482	0.71	0.78	0.844
σ1/σ0=5	ρ = 1	0.016	0.042	0.06	0.062	0.066
	ρ = 0.95	0.01	0.032	0.046	0.046	0.05
	ρ = 0	0.034	0.116	0.288	0.37	0.516
Panel C. γ= 10 (one level break)						
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.742	0.606	0.476	0.464	0.42
	ρ = 0.95	0.834	0.726	0.622	0.628	0.646
	ρ = 0	1	1	1	1	1
σ1/σ0=0.5	ρ = 1	0.988	0.958	0.896	0.836	0.794
	ρ = 0.95	0.992	0.98	0.962	0.938	0.922
	ρ = 0	1	1	1	1	1
σ1/σ0=3	ρ = 1	0.03	0.056	0.076	0.08	0.072
	ρ = 0.95	0.036	0.056	0.098	0.13	0.118
	ρ = 0	0.764	0.884	0.872	0.806	0.768
σ1/σ0=5	ρ = 1	0.016	0.036	0.068	0.068	0.064
	ρ = 0.95	0.014	0.042	0.044	0.064	0.064
	ρ = 0	0.302	0.612	0.812	0.818	0.826

# of Monte Carlo Simulations: 500
# of Bootstraps: 100
One vol break at: τσ=0.5
One level break at: τ=0.5

Table 3.B Finite sample powers of nominal 0.05-level tests: normal innovations with a single volatility break, WBS

Panel A. T=150, γ= 0 (No level break)

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.024	0.044	0.076	0.086	0.098
	ρ = 0.95	0.022	0.038	0.024	0.044	0.042
	ρ = 0	0.016	0.028	0.042	0.038	0.044
σ1/σ0=3	ρ = 1	0.018	0.052	0.066	0.074	0.078
	ρ = 0.95	0.014	0.044	0.056	0.066	0.062
	ρ = 0	0.008	0.02	0.056	0.05	0.042
σ1/σ0=5	ρ = 1	0.024	0.042	0.066	0.078	0.076
	ρ = 0.95	0.012	0.046	0.066	0.068	0.058
	ρ = 0	0.008	0.024	0.048	0.042	0.038

Panel B. T=150, γ= 5 (one level break)

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.08	0.11	0.096	0.112	0.126
	ρ = 0.95	0.11	0.114	0.13	0.19	0.232
	ρ = 0	0.966	0.984	0.964	0.946	0.9
σ1/σ0=0.5	ρ = 1	0.13	0.126	0.108	0.148	0.16
	ρ = 0.95	0.162	0.18	0.174	0.252	0.31
	ρ = 0	0.98	0.984	0.974	0.96	0.944
σ1/σ0=3	ρ = 1	0.016	0.054	0.076	0.086	0.086
	ρ = 0.95	0.014	0.04	0.05	0.078	0.068
	ρ = 0	0.228	0.562	0.742	0.776	0.798
σ1/σ0=5	ρ = 1	0.02	0.04	0.07	0.078	0.074
	ρ = 0.95	0.014	0.032	0.068	0.074	0.056
	ρ = 0	0.042	0.11	0.258	0.372	0.504

Panel C. T=150, γ= 10 (one level break)

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.742	0.606	0.476	0.464	0.42
	ρ = 0.95	0.834	0.726	0.622	0.628	0.646
	ρ = 0	1	1	1	1	1
σ1/σ0=0.5	ρ = 1	0.872	0.756	0.644	0.616	0.568
	ρ = 0.95	0.95	0.864	0.814	0.782	0.786
	ρ = 0	1	1	1	1	1
σ1/σ0=3	ρ = 1	0.038	0.068	0.09	0.112	0.134
	ρ = 0.95	0.044	0.096	0.104	0.15	0.182
	ρ = 0	0.844	0.9	0.862	0.798	0.728
σ1/σ0=5	ρ = 1	0.018	0.054	0.074	0.08	0.086
	ρ = 0.95	0.014	0.048	0.066	0.072	0.09
	ρ = 0	0.35	0.664	0.782	0.742	0.682

# of Monte Carlo Simulations: 500
# of Bootstraps: 100
One vol break at: τσ=0.75
One level break at: τ=0.5

Table 3.C Finite sample powers of nominal 0.05-level tests: normal innovations with a single volatility break, WBS

Panel A. T=150, γ= 0 (No level break)

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.018	0.038	0.06	0.076	0.092
	ρ = 0.95	0.01	0.028	0.036	0.048	0.058
	ρ = 0	0.004	0.018	0.03	0.022	0.03
σ1/σ0=3	ρ = 1	0.052	0.086	0.096	0.094	0.082
	ρ = 0.95	0.06	0.068	0.078	0.082	0.1
	ρ = 0	0.034	0.068	0.068	0.07	0.062
σ1/σ0=5	ρ = 1	0.068	0.088	0.094	0.098	0.098
	ρ = 0.95	0.074	0.076	0.078	0.086	0.096
	ρ = 0	0.034	0.064	0.06	0.054	0.034

Panel B. T=150, γ= 5 (one level break)

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.08	0.11	0.096	0.112	0.126
	ρ = 0.95	0.11	0.114	0.13	0.19	0.232
	ρ = 0	0.966	0.984	0.964	0.946	0.9
σ1/σ0=0.5	ρ = 1	0.086	0.108	0.108	0.124	0.14
	ρ = 0.95	0.124	0.14	0.152	0.208	0.26
	ρ = 0	0.944	0.962	0.94	0.916	0.888
σ1/σ0=3	ρ = 1	0.054	0.086	0.092	0.1	0.092
	ρ = 0.95	0.056	0.072	0.088	0.1	0.096
	ρ = 0	0.706	0.926	0.958	0.966	0.73
σ1/σ0=5	ρ = 1	0.07	0.086	0.094	0.11	0.112
	ρ = 0.95	0.08	0.082	0.074	0.088	0.084
	ρ = 0	0.102	0.394	0.62	0.708	0.554

Panel C. T=150, γ= 10 (one level break)

		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.742	0.606	0.476	0.464	0.42
	ρ = 0.95	0.834	0.726	0.622	0.628	0.646
	ρ = 0	1	1	1	1	1
σ1/σ0=0.5	ρ = 1	0.736	0.626	0.524	0.49	0.458
	ρ = 0.95	0.842	0.77	0.68	0.698	0.694
	ρ = 0	1	1	1	1	1
σ1/σ0=3	ρ = 1	0.122	0.128	0.126	0.132	0.138
	ρ = 0.95	0.166	0.188	0.21	0.224	0.228
	ρ = 0	1	1	1	1	0.91
σ1/σ0=5	ρ = 1	0.07	0.092	0.096	0.112	0.102
	ρ = 0.95	0.086	0.092	0.096	0.116	0.106
	ρ = 0	0.896	0.996	1	1	0.72

# of Monte Carlo Simulations: 500
# of Bootstraps: 100
one vol break at: τσ=0.5
Three level breaks at: τ=[0.25,0.5,0.75]

Table 3.D Finite sample powers of nominal 0.05-level tests: normal innovations with a multiple volatility breaks, WBS

		U				
Panel A. T=150, γi= 0 (No level break)						
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.012	0.038	0.05	0.054	0.068
	ρ = 0.95	0.006	0.03	0.046	0.05	0.062
	ρ = 0	0.006	0.026	0.024	0.026	0.034
σ1/σ0=0.5	ρ = 1	0.024	0.044	0.076	0.086	0.098
	ρ = 0.95	0.022	0.038	0.024	0.044	0.042
	ρ = 0	0.016	0.028	0.042	0.038	0.044
σ1/σ0=3	ρ = 1	0.018	0.052	0.066	0.074	0.078
	ρ = 0.95	0.014	0.044	0.056	0.066	0.062
	ρ = 0	0.008	0.02	0.056	0.05	0.042
σ1/σ0=5	ρ = 1	0.024	0.042	0.066	0.078	0.076
	ρ = 0.95	0.012	0.046	0.066	0.068	0.058
	ρ = 0	0.008	0.024	0.048	0.042	0.038
Panel B. T=150, γi= 5 (three level breaks)						
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.174	0.13	0.154	0.19	0.064
	ρ = 0.95	0.248	0.178	0.3	0.386	0.108
	ρ = 0	0.946	0.978	0.966	0.952	0.166
σ1/σ0=0.5	ρ = 1	0.208	0.154	0.182	0.236	0.068
	ρ = 0.95	0.324	0.252	0.332	0.43	0.058
	ρ = 0	0.974	0.978	0.98	0.98	0.06
σ1/σ0=3	ρ = 1	0.026	0.052	0.064	0.084	0.07
	ρ = 0.95	0.026	0.058	0.08	0.09	0.086
	ρ = 0	0.284	0.534	0.596	0.554	0.392
σ1/σ0=5	ρ = 1	0.024	0.05	0.07	0.078	0.078
	ρ = 0.95	0.012	0.044	0.068	0.078	0.068
	ρ = 0	0.076	0.192	0.35	0.402	0.402
Panel C. T=150, γi= 10 (three level breaks)						
		U				
		m = 0.10	m = 0.15	m = 0.20	m = 0.25	m = 0.30
constant volatility	ρ = 1	0.908	0.808	0.704	0.696	0.052
	ρ = 0.95	0.972	0.93	0.884	0.892	0.022
	ρ = 0	1	1	1	1	0.01
σ1/σ0=0.5	ρ = 1	0.96	0.894	0.834	0.8	0.026
	ρ = 0.95	0.992	0.962	0.956	0.95	0.002
	ρ = 0	1	1	1	1	0
σ1/σ0=3	ρ = 1	0.066	0.064	0.098	0.148	0.05
	ρ = 0.95	0.09	0.118	0.19	0.27	0.12
	ρ = 0	0.712	0.83	0.782	0.744	0.114
σ1/σ0=5	ρ = 1	0.03	0.052	0.062	0.082	0.052
	ρ = 0.95	0.034	0.062	0.084	0.104	0.09
	ρ = 0	0.374	0.62	0.612	0.54	0.22

Chapter 5

Statistical arbitrage with stochastic cointegration

This chapter provides empirical evidence supporting the use of relative pricing in pairs of financial securities that exhibit profound statistical relationship. As a means to formally detect pairs of stocks that co-moved historically in the equity market, we propose to use the new nonlinear method for cointegration analysis derived by Harris, McCabe and Leybourne (2002, 2006). We test for the presence of stochastic cointegration and empirically assess the nonlinear generalization of a new paradigm over the standard method of Engle and Granger (1987). Compared with the standard method, stochastic cointegration is found to be an effective tool that consistently estimates long-run states of equilibria and repeatedly detects such relationships in the US equity market.

5.1 Introduction

Pair Trading is an investment strategy that has long been employed by hedge funds and proprietary trading desks. In its most common form, the concept of the strategy is simple: find two financial securities that have moved together historically. When the spread between them widens significantly, (short) sell the winner and (long) buy the loser. As soon as the spread reverts back to its 'norm' the position in the pair is closed, resulting in net gain. The profitability of pairs trading strategies relies heavily on the expectation that the relationship observed between the securities in the past will prevail in the future. Any significant deviations from the modeled relationship are assumed to be temporary and are expected to revert. Questions such as how to define and identify the long-run 'norm' between prices of securities, as well as what constitutes significant divergence from the long-run state all need to be formalized and modeled directly.

A well documented example that offers a theoretical explanation for why prices of different securities tend to move together can be found in the stock market. Arbitrage Pricing Theory (APT) developed primarily by Ross (1976a) and Ross (1976b), suggests that the price of a share in a company should equal the net present value of the sum of its future dividends. The discount rate that is used to discount future dividends of stocks is particularly sensitive to a shift in the expectation of future interest rate. Ceteris paribus, an increase in the expected interest rate generally decreases the net present value of stocks. Conversely, an expected decrease in future interest rate usually results in an increase in stock prices. Stock prices therefore, tend to co-move in response to changes in macroeconomic variables. Other economic news often has a more pronounced effect on prices of a subset of stocks, stocks that serve as close economic substitutes to each other. For example, a sharp unexpected price increase of a fundamental production component such as crude oil is likely to have a substantial negative effect on the profitability of stocks that belong to the airline industry. In response to a price increase of a common production factor (oil), stock prices of the airline industry have high propensity to co-move in the same direction to a lower level.

The Law of One Price as defined by Ingersoll (1987) maintains that securities that have the *exact* same payoffs in every state of the world must sell for the *exact* same price. Any discrepancies between prices of identical securities usually disappear almost as soon as they appear as a result of (riskless) arbitrage activity. A pure arbitrageur is engaged in the activity of buying the underpriced security and simultaneously selling the same security in the market where it is overpriced. This activity takes place until the price discrepancy disappears. Occasionally, (slightly) different prices can be observed for the same stock that is traded simultaneously on two different exchanges. The arbitrageur buys the stock on the exchange where it is underpriced and at the same time sells it on the exchange where it is overpriced, netting a riskless gain. The concept of pure arbitrage can be extended to any securities that maintain an exact arbitrage relationship or any securities whose future payoff can be perfectly mimicked using a combination of other financial instruments, such as derivatives. Chen and Knez (1995) extend the definition of Ingersoll and argue that *closely integrated* securities should have *similar* prices. This weaker condition on the prices is the main motivation for strategies of statistical arbitrage, where prices of similar securities are expected to be in the same neighborhood, in a statistical sense, but not exactly the same, as in the case of pure arbitrage.

Trading strategies of statistical arbitrage aim to exploit deviations from a statistical relationship that is observed between securities. Similar to Jegadeesh and Titman (1993), Gatev, W.N.Goetzmann and Rouwenhorst (2006) test the assumption that there is a potential profitability from pairs trading simply due to the assumption that there is a tendency of stock prices to revert to their means at certain horizons. To address this hypothesis, they develop a bootstrapping test based upon random pair choice; if pairs trading profits were simply due to mean-reversion, then randomly chosen pairs should generate profits by buying loser and selling winner stocks. This simple contrarian strategy was found to be unprofitable over the period of their study, suggesting that the mean-reversion assumption by itself does not tell the whole story. They propose to use cointegration analysis instead.

Alexander (1999) shows that high correlation of returns does not necessarily imply cointegration in prices. Correlation measures co-movements in returns, which are subject to great instabilities over time. It measures the degree of linear association between a set of variables but does not carry information on the long-term relationship that may exist between the variables. Hence, trading and hedging strategies that are based on correlation require frequent parameter estimation and rebalancing. Cointegration on the other hand, measures long time co-movement in prices, which prevails even through periods when static correlation appears low. Investment management and trading strategies that are based only on volatility and correlation of returns cannot guarantee long term performance. Cointegration extends the traditional models by including a preliminary stage in which the multivariate price data are analyzed, and then augments the correlation analysis by including the dynamics and causal flows between returns. Mispricing and over-hedging may occur if cointegration is ignored. Cointegration in financial assets can be found between spot and futures, bonds of different maturities, bonds issued in different counties, international indices or in fact anywhere where spreads are mean-reverting.

Bossaerts and Green (1989), as well as Gatev et al. (2006) suggest that pairs trading strategies may be justified within an equilibrium asset-pricing framework with non stationary common factors. They employ cointegration analysis as a tool to identify stocks whose prices move together over a given history. Originally proposed by Engle and Granger (1987) (EG), the methodology provides a framework in which the long and short components of a pair can fluctuate around a nonstationary factor, allowing for a long-run state of equilibrium to prevail between the stocks. Using the standard EG framework, Bossaerts (1988) and Gatev et al. (2006) find evidence of price cointegration in the US equity market.

The EG methodology indeed has had some degree of success in detecting cointegrating pairs in the equity market, but the procedure provides little support to an important economic theory, the term structure of interest rates. The theory suggests that short and long term rates should cointegrate with a single cointegrating vector equal to one. In their empirical analysis, Campbell and Shiller (1987) did not find cointegration between the rates and argue that the spreads between them tend to "move-to-much" to be consistent with the EG paradigm. Harris et al. (2002) (HML) propose a new framework for cointegration, one that allows for volatility in excess of that catered for by the standard integration/cointegration paradigm through the introduction of nonstationary heteroscedasticity. Using this technique they are able to provide empirical support to the term structure theory and a cointegrating relationship between the rates is indeed detected. HML offers a nonlinear generalization over the standard EG paradigm.

Similar to HML, Xiao (2009) proposes an alternative way to model the nonlinearity in cointegrated systems observed in the equity market via a quantile cointegration model. His model settings allow the cointegrating coefficients to be non-constant, that is to evolve over time. Sollis (2008) favors the heteroskedastic integration (HI)/stochastic cointegration (SC) framework and points out that the HML heteroscedastically integrated processes can be represented as models with time-varying parameters, a property that better captures the dynamics seen in financial and economic data.

In this chapter we maintain that while cointegration analysis is a relevant tool for identifying stocks that co-move over time, the EG framework is sub-optimal with respect to both the quality and the number of detected cointegrating pairs. Instead, the SC framework of HML provides a superior alternative. The advantage of employing SC over the standard EG is threefold: 1) Using the HML method, a larger number of cointegrating pairs is detected, providing the statistical arbitrageur with more opportunities to take advantage of price discrepancies. 2) Unlike EG, the long-run parameters estimated by the HML technique are consistent in the presence of heteroscedastically integrated regressors. 3) The HML method identifies the specific form of cointegration (heteroscedastic or stationary), allowing the statistical arbitrageur to further fine-tune their trading models to take advantage of the type of the long-run relationship.

The chapter is structured as follows: Section 2 introduces the stochastic integration/cointegration paradigm and highlights the differences between the new approach and the standard EG one. Section 3 assesses empirically the value of the new framework over the standard one by comparing the performance of the two approaches in detecting cointegrating relationships in the US equity market. The results validate that relative pricing is an effective pricing methodology when applied to a specific set of stocks, stocks that are stochastically cointegrated. Section 4 concludes.

5.2 Stochastic Integration

A popular way to model time series of prices of stocks is to use unit root models such as random walk. These however, fail to capture an important stylized fact that is observed in many financial data in the form of excess volatility. Conditional heteroscedasticity, particularly in a form that allows level dependent one, is often a desired feature in modeling financial data. The specification of HML offers an adequate alternative as it allows for level dependent heteroscedasticity while assuming prices are stochastically trending. Their concept builds on the heteroscedastically integrated processes that were originally studied by Hansen (1992). For the purpose of modeling a system of stock prices we adopt the model setting as proposed in McCabe et al. (2006):

$$\begin{aligned} Z_t &= \mu + \delta_t + \Pi w_t + \varepsilon_t + V_t h_t \qquad (5.1) \\ w_t &= w_{t-1} + \eta_t \\ h_t &= h_{t-1} + \nu_t \end{aligned}$$

for $t = 1, \ldots, T$. Where Z_t, μ, δ and ε_t are $m \times 1$ vectors; w_t and η_t are $n \times 1$ vectors; h_t and ν_t are $p \times 1$ vectors; Π and V_t are $m \times n$ and $m \times p$ matrices, respectively. The disturbances $\varepsilon_t, \eta_t, \nu_t$ and V_t are mean zero stationary processes. In this case Z_t represents a vector of time series of observed stock prices, consists of the deterministics μ and δ, an integrated component Πw_t, and a shock term $\varepsilon_t + V_t h_t$. The shock term has a linear component, ε_t, and a nonlinear component, $V_t h_t$, that is nonstationary heteroscedastic from its dependence on the $I(1)$ process h_t. When h_t is replaced with w_t the process exhibits level dependent heteroscedastic behavior. The Linear Process (LP) assumption is made for the statistical properties of the disturbance terms in equation (5.1), allowing for general forms of serial correlation, cross-correlation and endogeneity.

Assumption 5.1 (LP). *Let* $\zeta = [\nu_t', vec(V_t)', \eta_t', \varepsilon_t']'$ *be generated by the vector linear process* $\zeta_t = \sum_{j=0}^{\infty} C_j \xi_{t-j}$ *where*

(a) $\sum_{j=0}^{\infty} j||C_j|| < \infty$ *with* C_0 *having full rank*[1].

(b) ξ_t *is an i.i.d. sequence.*

(c) $E(\xi_t \xi_t') = I$.

(d) *For all* $E(\xi_{it}^{16})$ *is bounded.*

From equation (5.1) HML derive the process for the individual elements of Z_t. Let e_i be the $m \times 1$ vector with 1 in its i^{th} position and 0 elsewhere, so that $e_i' Z_t = z_{it}$, the i^{th} element of the vector Z_t, such that

$$z_{it} = e_i' \Pi w_t + e_i'(\varepsilon_t + V_t h_t)$$

where we make the simplifying assumption that $\mu = \delta = 0$. If $e_i'\Pi \neq 0$ then z_{it} is said to be *stochastically integrated* (SI). If in addition, $e_i' E(V_t V_t') e_i > 0$ then z_{it} is said to be *heteroscedastically integrated* (HI) due to the term $e_i' V_t h_t$. On the other hand if $e_i' V_t = 0$ then z_{it} has a constant unit root, that is simply $I(1)$. A stochastically integrated variable therefore, nests both forms of integration: constant unit root and the heteroscedastic one.

5.2.1 Stochastic Cointegration

We aim to model the linear relationships between the time series of stock prices in Z_t. This can be achieved as follows: let c be a non-zero $m \times 1$ vector and

[1] $||A|| = \sqrt{tr(AA')}$.

consider

$$c'Z_t = c'\Pi w_t + c'(\varepsilon_t + V_t h_t)$$

When $c'\Pi = 0$, the variables in Z_t are said to be *stochastically cointegrated*, otherwise the variables are not cointegrated. HML note that under stochastic cointegration $c'Z_t = c'(\varepsilon_t + V_t h_t)$ behaves like a stochastically integrated process net of its stochastic trend component and they refer to such a process as being *stochastically trendless*. Under Assumption LP, they show that as $s \to \infty$ (with t fixed),

$$E(\varepsilon_{t+s} + V_{t+s}h_{t+s} \mid \mathfrak{I}_t) - E(\varepsilon_{t+s} + V_{t+s}h_{t+s}) \to 0$$

That means that the behavior of the process up to time t a has negligible effect on its behavior into the infinite future. Therefore, even though the disturbances ν_t have an infinitely persistent effect on h_{t+s} their effect on the *level* of $V_{t+s}h_{t+s}$ is only transitory. This in turn, implies that the product process $V_t h_t$ is stochastically trendless, even if V_t is correlated with ν_t. Despite the fact that $V_t h_t$ is nonstationary heteroscedastic (as it exhibits linearly trending variance), it is the stochastically trendless nature of $c'Z_t = c'(\varepsilon_t + V_t h_t)$ that facilitates co-movement of a nonstationary heteroscedastic type. The process described in equation (5.1) is a departure from the standard cointegration framework of EG, as EG assume the cointegrating residuals are asymptotically stationary, while HML require them to be only stochastically trendless.

Similar to the interpretation of z_{it} above, when $c'\Pi = 0$ and $c'E(V_t V_t')c = 0$,

$c'Z_t = c'\varepsilon_t$ is stationary. If in addition, $V_t = 0$ then the variables are integrated and cointegrated in the standard EG sense. Because of the stationary behavior of $c'Z_t$ in either case, HML refer to this as *stationary cointegration*. When $c'\Pi = 0$ and $c'E(V_tV_t')c > 0$ the variables Z_t are said to be *heteroscedastically cointegrated* (HC). Thus, stochastic cointegration encompasses both stationary cointegration (possibly of the EG kind) and heteroscedastic cointegration. HML essentially replace the restrictive requirement of EG on the cointegrating errors with a weaker condition. The concept of SC is weaker than the conventional EG paradigm as it only requires the residuals from the long-run regression not to be $I(1)$, rather than requiring them to be $I(0)$ stationary. While EG impose $I(0)$ stationary behavior on the residuals of the long-run model, HML only require that $I(1)$ behavior is absent from them.

5.2.2 Hypothesis Tests and Test Statistics

A formal statistical test needs to be employed in order to identify pairs of stocks that co-moved historically in a stochastically cointegrated fashion. HML develop a procedure that tests whether the system is cointegrated and set the null of the test as a cointegrated system and the alternative as a non-cointegrated system: $H^0 : c'\Pi = 0$ and $H^1 : c'\Pi \neq 0$. Further, within stochastic cointegration, they develop a procedure that tests whether the cointegration is stationary or of heteroscedastic form. The null of stationary cointegration against the heteroscedastic alternative is tested by partitioning H^0 as

$$H_0^0 : \quad c'E(V_tV_t')c = 0 \text{ and } H_1^0 : \quad c'E(V_tV_t')c > 0$$

The system described in equation (5.1) provides merely a representation of the model and cannot be estimated directly since the only observed variables are those in Z_t. In order to estimate the system HML partition Z_t in equation (5.1 into a scalar y_t and an $(m-1)\times 1$ vector x_t as $Z_t = [y_t, x_t']'$ so that it can be estimated via a regression equation:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix} + \begin{bmatrix} \delta_y \\ \delta_x \end{bmatrix} t + \begin{bmatrix} \pi_y' \\ \Pi_x \end{bmatrix} w_t + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{bmatrix} + \begin{bmatrix} v_{yt}' \\ V_{xt} \end{bmatrix} h_t \tag{5.2}$$

Where y_t, μ_y, δ_y and ε_{yt} are scalars, x_t, μ_x, δ_x and ε_{xt} are $(m-1)\times 1$ vectors, π_y and v_{yt} are $n\times 1$ and $p\times 1$ vectors, respectively, while Π_x and V_{xt} are $(m-1)\times n$ and $(m-1)\times p$ matrices. Letting $c = [1, -\beta']'$, $\alpha = \mu_y - \beta'\mu_x$, $k = \delta_y - \beta'\delta_x$, $e_t = \varepsilon_{yt} - \beta'\varepsilon_{xt}$, $q' = \pi_y' - \beta'\Pi_x = c'\Pi$ and $v_t' = v_{yt}' - \beta' V_{xt} = c'V_t$, then we have

$$y_t = \alpha + kt + x_t'\beta + u_t \tag{5.3}$$

$$u_t = e_t + q'w_t + v_t'h_t \tag{5.4}$$

The regression error term u_t is the equilibrium residual from the long-run equation. It has a stationary term e_t, the integrated term $q'w_t$ and the heteroscedastic component $v_t'h_t$[2]. It is assumed there is only one cointegrating vector so that $rank(\Pi_x) = m-1$, which imposes the restriction that $n \geq m-1$. This means that any further sub-relationships among the x_t variables in Eq. (5.3) are excluded. The null hypothesis of stochastic cointe-

[2] u_t need not have zero mean so that α is not an intercept in the usual sense.

gration against alternative of non-cointegration is expressed via Eq. (5.3) as $H^0 : q = 0$ and $H^1 : q \neq 0$. Within the cointegrating relationship, H^0, the null hypothesis of stationary cointegration against the heteroscedastic alternative is $H_0^0 : E(v_t'v_t) = 0$ against $H_1^0 : E(v_t'v_t > 0)$.

Both tests, H^0 vs. H^1 and H_0^0 vs. H_1^0, use lag covariances and long-run variances as inputs. Lag covariances for an arbitrary process $\{a_t\}$ is defined by $\gamma_j(a_t) = T^{-1}\sum_{s=j+1}^{T} a_s a_{s-j}$. A HAC (heteroscedasic and autocorrelation consistent) estimator of the long run variance (LRV) is defined by

$$\omega^2(a_t) = \gamma_0(a_t) + 2\sum j = 1^l \lambda(j/l)\gamma_j(a_t) \tag{5.5}$$

where $\lambda(\cdot)$ is a window with lag truncation parameter l. It is assumed that Assumption KN below holds.

Assumption 5.2 (KN, *(Kernel and lag Length)*). *(a)* $\lambda(0) = 1$

(b) $0 \leq \lambda(x) \leq 1$ *for* $0 \leq x < 1$

(c) $\lambda(x)$ *is continuous and of bounded variation on* $[0, 1]$.

(d) $l \to \infty$ *as* $T \to \infty$

Testing for stochastic cointegration against non-cointegration (testing H^0 against H^1) is equivalent to testing whether $q = 0$ in $u_t = e_t + q'w_t + v_t'h_t$. The null encompasses both stationary and heteroscedastic cointegration; while the alternative is $I(1)$ or heteroscedastic integration. As an optimal test HML consider the statistic

$$S_{nc} = \sum_{t=k+1}^{T} u_t u_{t-k} \tag{5.6}$$

When all the disturbances are *i.i.d.* S_{nc} with $k = 1$ basically tests for zero autocorrelation in u_t against the correlation induced by the $I(1)$ term $q'w_t$. By contrast, when the disturbances terms are more general than *i.i.d.*, S_{nc} needs to be modified to eliminate the nuisance parameters that result from the autocorrelation and from the presence of $v_t'h_t$. A solution is obtained by allowing k to increase with T. HML show that under the cointegration null, H^0, the statistic S_{nc} is asymptotically $N(0,1)$[3] and is consistent under the alternative of no cointegration, H^1. Letting k become large eliminates any correlation between u_t and u_{t-k}.

Since y_t and x_t are observed, $b = [\alpha, k, \beta']'$ of Eq (5.3) is estimated by means of the estimator $\hat{b}_k = [\hat{\alpha}_k, \hat{k}_k, \hat{\beta}_k']'$ given by

$$\hat{b}_k = \left(\sum_{t=k+1}^{T} X_{t-k} X_t \right)^{-1} \sum_{t=k+1}^{T} X_{t-k} y_t \tag{5.7}$$

where $X_t = [1, t, x_t']'$. This estimator is called an Asymptotic IV (AIV).

As opposed to AIV estimator, the OLS estimator that is used in the EG framework is not consistent under heteroscedastic cointegration unless consists entirely of an $I(1)$ process. In this case, any economic decision (such as trading strategies, risk management, hedging, etc.) that is based on the long-run equilibrium estimate obtained via the EG estimation procedure may not be

[3]when standardized with a HAC estimator.

reliable.

Eq (5.6) is now reconstructed using the AIV residuals:

$$\hat{u}_t = y_t - \hat{\alpha}_k - \hat{k}_k t - x_t' \hat{\beta}_k. \tag{5.8}$$

Using limit theory HML prove consistency of the AIV estimator and asymptotic normality subject to some additional exogeneity restrictions. They show that under H^0 the test statistic and its distribution is:

$$\hat{S}_{nc} = \frac{T^{-1/2} \sum_{t=k+1}^{T} \hat{u}_t \hat{u}_{t-k}}{\sqrt{\omega^2(\hat{u}_t \hat{u}_{t-k})}} \xrightarrow{d} N(0,1)$$

while under H^1 the distribution of $|\hat{S}_{nc}|$ diverges[4] as $T \to \infty$.

Testing H_0^0 against H_1^0

In decomposing the composite hypothesis H^0 into null of stationary cointegration against heteroscedastic alternative, we need to test whether $E(v_t' v_t) = 0$ in 5.4, maintaining $q = 0$. McCabe and Leybourne (2000) show that a locally most powerful test of H_0^0 against H_1^0 is given by

$$S_{hc} = \sum_{t=1}^{T} t u_t^2 \tag{5.9}$$

[4] $\hat{u}_t$ is defined in (5.8) using (5.7); $\omega^2()$ is defined in (5.5).

HML show that under H_0^0 the test statistic is normally distributed and the statistic is:

$$\hat{S}_{hc} = (1/12)^{1/2} \frac{T^{-3/2} \sum_{t=1}^{T} t(\hat{u}_t^2 - \hat{\sigma}_u^2)}{\sqrt{\omega^2(\hat{u}_t^2 - \hat{\sigma}_u^2)}} \xrightarrow{d} N(0,1)$$

while under H_1^0, the distribution of $|\hat{S}_{hc}|$ diverges as $T \to \infty$.

$\hat{S}_{hc}$ is calculated using $\hat{u}_t^2 - \hat{\sigma}_u^2$ rather than simply $\hat{u}_t^2$ as 5.9 might suggest. This alteration is needed to center the statistic and make sure it is invariant to the variance of u_t under H_0^0. The structure of S_{hc} can also be used to test the null of $I(1)$ against the alternative of *HI* for any given individual series, by simply constructing $\hat{S}_{hc}$ by redefining $\hat{u}_t$ as $\hat{u}_t = \Delta y_t - \hat{\delta}_y$ where $\hat{\delta}_y$ is an estimator of the trend coefficient δ_y given by $\hat{\delta}_y = T^{-1} \sum_{t=1}^{T} \Delta y_t$. HML denote this statistic $\hat{S}_{hi}$ and show that $\hat{S}_{hi} \xrightarrow{d} N(0,1)$ if y_t is $I(1)$ and $|\hat{S}_{hi}|$ diverges if y_t is *HI*. The same conclusion arise if linear trends are excluded from (5.3), in which case $\hat{u}_t = \Delta y_t$.

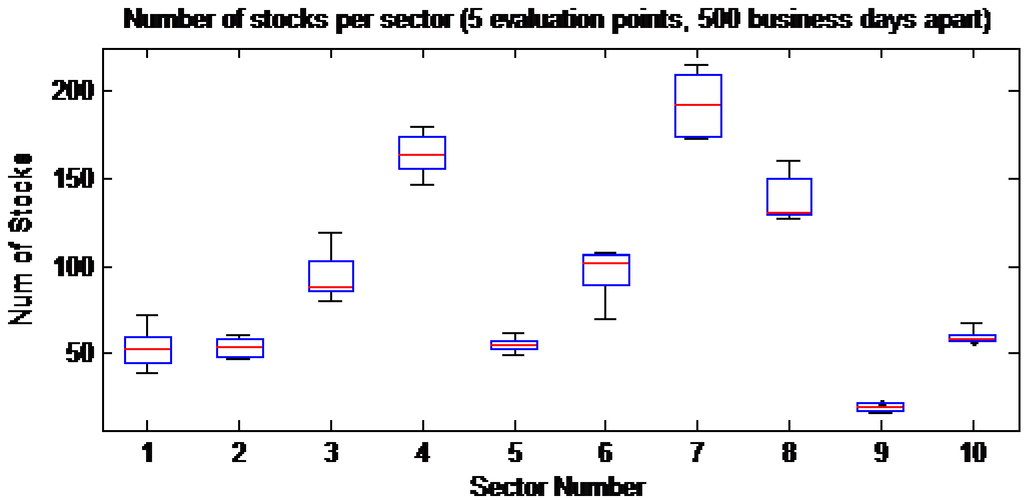
5.3 Empirical Methodology and Results

The objective of this section is twofold. First, we provide empirical evidence suggesting that *stochastic integration* is evident in time series of daily stock prices. Second, we show that the framework of *stochastic cointegration* can be used to effectively detect pairs of stocks whose prices co-moved historically. Gatev et al. (2006) form pairs based on an algorithm that minimizes the distance between (standardized) historical prices of two stocks. This algorithm is indeed appealing but it excludes other forms of potentially robust relationships between two stocks. Cointegration analysis does not necessarily form pairs with minimal distance between the variables, but clearly has the ability to detect stocks that co-moved in a well defined fashion. Within the cointegration framework, cointegrating pairs of stocks that exhibit minimal distance are likely to have a long-run regression slope close to one, so that they closely follow each other with minimal tracking error. We use instead, the cointegrating technique as a pairs forming algorithm as it is likely to produce a larger universe of pairs; pairs that co-move with cointegrating vector not necessarily $[1, -1]$.

We analyze prices of stocks that belong to the Russell 1000 index. The index consists of the 1000 largest firms in the US. The index is updated annually; some stocks are deleted from the index because they no longer meet the index membership criteria while others are added. The dataset consists of daily closing prices of the index stock members, going back about 10 years in time. We group the stocks into 10 sectors, as classified by Global Industry Classification Standard[5]. In order to test for cointegration and highlight some characteristics of the data we arbitrarily divide the data into 5 consecutive, non-overlapping time sections, each time section spans over two years and consists of 500 observations (there are about 250 business days in a calendar year).

[5]See the appendix for a full list of all the sectors.

Figure 5.1:



Time Section 1	yr:1-2	Observation: 1-500	08/Sep/1997 - 02/Sep/1999
Time Section 2	yr:3-4	Observation: 501-1000	03/Sep/1999 - 27/Aug/2001
Time Section 3	yr:5-6	Observation: 1001-1500	28/Aug/2001 - 27/Aug/2003
Time Section 4	yr:7-8	Observation: 1501-2000	28/Aug/2003 - 22/Aug/2005
Time Section 5	yr:9-10	Observation: 2001-2500	23/Aug/2005 - 17/Aug/2007

At the end of each of the 5 time sections we calculate the number of stocks in each of the 10 groups, counting only stocks that have full history of daily prices during the respective time section (500 observations[6]). Fig.1 shows the variation in the number of stocks in each sector across time, using a boxplot[7].

[6]We conducted the analysis using 250 observations too and the qualitative conclusion was not changed.

[7]A Boxplot produces a box and whisker plot for each column of a matrix. The box has lines at the lower quartile, median, and upper quartile values. Whiskers extend from each end of the box to the adjacent values in the data, the most extreme values within 1.5 times the interquartile range from the ends of the box. Outliers are data with values beyond the ends of the whiskers. Outliers are displayed with a red + sign.

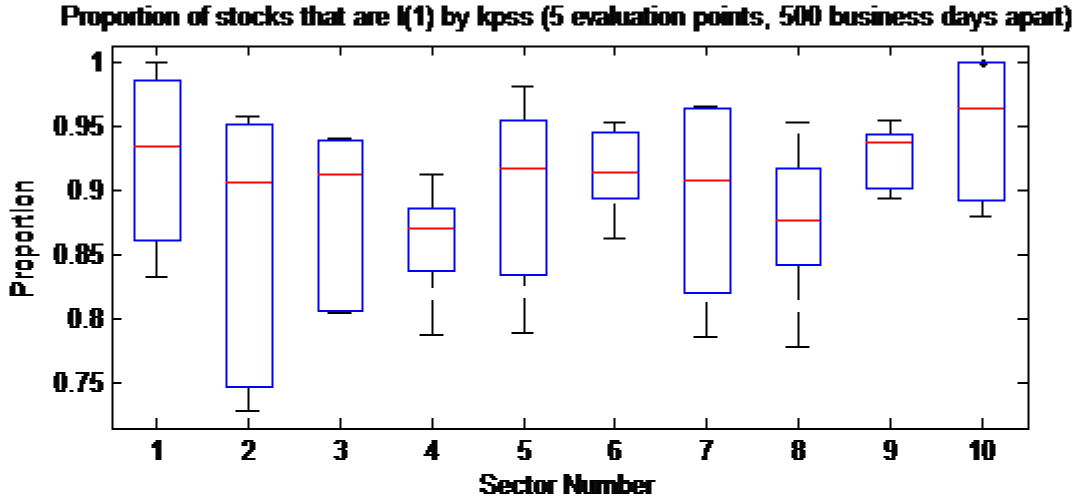
There are 5 observations in each box, corresponding to the 5 sequential time sections. For example, the number of stocks that belong to Sector 1 (Energy and consumption) at each of the 5 time sections is: 39, 53, 46, 55, and 72. The median number of stocks across time for this sector is 53, represented in Fig.1 by the red horizontal line. The figure also shows that Sector 7 (Financials) has the highest median number of stocks in a group over time, 192, while Sector 9 (Telecommunication Services) has the lowest median, 19. Overall, the number of stocks per sector seems rather stable across time.

As mentioned above, it is common practice to assume that stock prices follow I(1) process. In order to test the I(1) hypothesis, at the end of each time section we calculate for each sector the proportion of stocks that are identified as I(1) using the KPSS test (Kwiatkowski, Phillips, Schmidt and Shin 1992), denoted K_s, for stationarity[8]. The null hypothesis for the KPSS test is I(0), and the alternative is I(1). Fig.2 shows the proportion of stocks that were found to be I(1) in each of the 5 time sections. Visually comparing the lengths of the 10 boxes, the figure suggests that Sector 2 exhibits the largest variation in the proportions of I(1) stocks across the five time sections, with as low as 73% of the stocks being I(1) at one time section and as high as 96% in other time section. Overall, based on K_s the median proportion of stocks that are I(1) exceed 85% (red horizontal lines).

Fig.2 merely confirms a well known stylized fact that daily stock prices are in general nonstationary. The next figure, however, will highlight the fact that a decent proportion of these I(1) stocks are actually better modeled as

[8]As the sample size at each time section is 500 observations, asymptotic values are used in the KPSS test. The test is performed with a constant and no trend.

Figure 5.2:



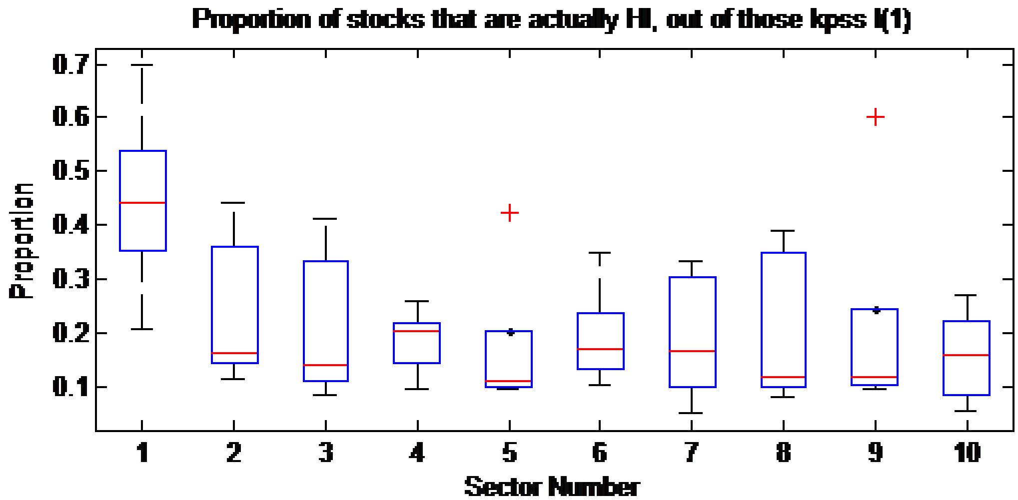
heteroscedastically integrated processes.

In each sector group we focus on the stocks that were found to be I(1) by KPSS and test each stock individually whether they are indeed I(1) or actually HI, using the S_{hi} test as detailed in Section 5.2.2.

Fig.3 shows that across time, at least 10% of the stocks that found to be I(1) by KPSS are better modeled as HI processes, as indicated by the median proportion in each sector group (the red horizontal lines). In Sector 1 for example, the proportion of stocks that were found to be HI reaches 70% at one time section and drops to about 20% in other time section. The figure points out that heteroscedastic integration is evident in the daily stock prices of the US market and should not be ignored if accurate modeling is desired.

Using the framework of stochastic integration and cointegration is particularly important when one searches for cointegrating pairs of stocks. In each of the 5 time sections, we search for cointegrating pairs among all possible

Figure 5.3:

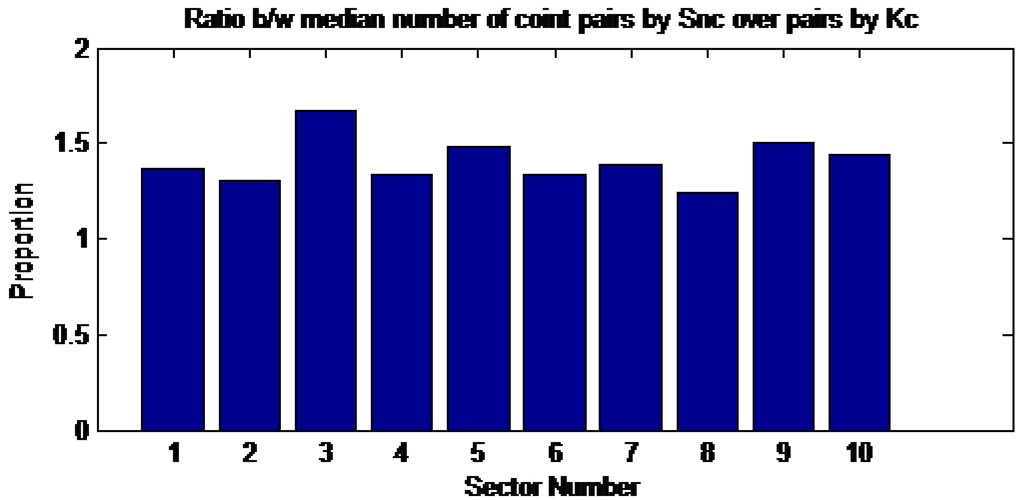


pair combinations within each sector. In this analysis a pair is identified as cointegrating if both stocks in a potential pair are I(1), as determined by the KPSS test for stationarity, and the residual-based test for stochastic cointegration S_{nc} indicates cointegration. The number of pairs that were found to be stochastically cointegrated is then recorded for each sector at each time section. The S_{nc} test is performed with a constant and no trend[9]. As in HML, for the AIV estimator we set $k = \lfloor T^{1/2} \rfloor$, ($\lfloor \cdot \rfloor$ denoting the integer part of). For the variance estimators we use Bartlett kernel for $\lambda(\cdot)$ and set the lag truncation parameter $l = 12(T/100)^{1/4}$ throughout.

Similarly, we search for cointegrating pairs under the EG assumptions using the residual test of Shin (1994), denoted K_c, testing for the null hypothesis of cointegration between I(1) series against the alternative of no cointegration.

[9] There are $n(n-1)$ possible combinations of pairs, where n is the number of stocks in the sector group at a certain time section. For example, in the 5th time section, Sector 1 has 72 stocks; this produces 5112 potential pairs. In this case, regressing stock A on stock B is treated and counted separately from regression B on A.

Figure 5.4:

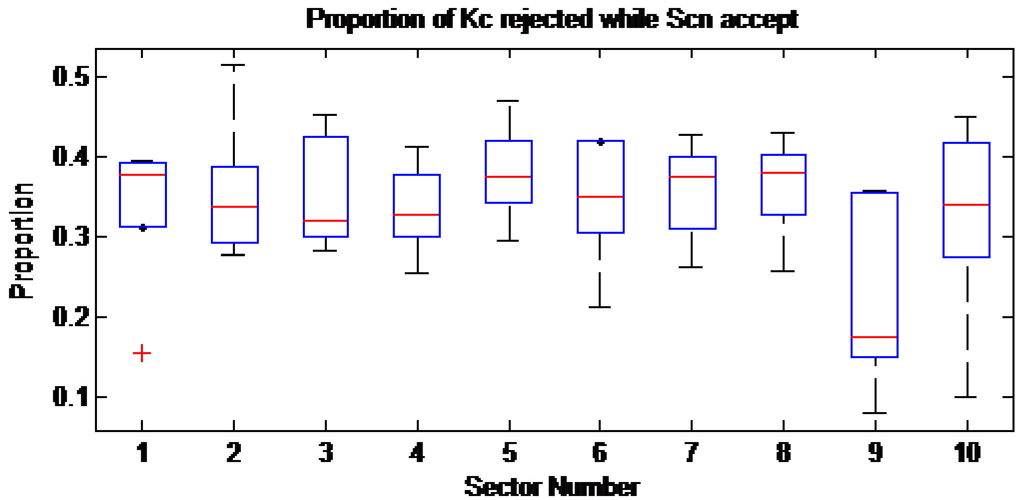


We use an efficient OLS estimator in which $\lfloor T^{1/4} \rfloor$ lead and lag terms in Δx_t are added into the regression equation of y_t on x_t (see Saikkonen (1991) for details). The test is performed with a constant and no trend. We then compare the number of pairs detected by the two methods.

Fig.4 depicts the median number of pairs detected by S_{nc} as a proportion of the median number of pairs that are detected by K_c. The figure shows that in each and every sector the median number of pairs detected by stochastic cointegration is significantly larger than that detected by the conventional method, by a factor of as low as 20% more pairs in Sector 8 and as high as 67% more pairs in Sector 3.

In order to further highlight the superiority of SC over the conventional Engle-Granger in detecting cointegrating pairs in the stock market, we perform the KPSS test for cointegration on each pair that is found to be stochastically cointegrated based on S_{nc}. The number of cointegrating pairs that are rejected

Figure 5.5:

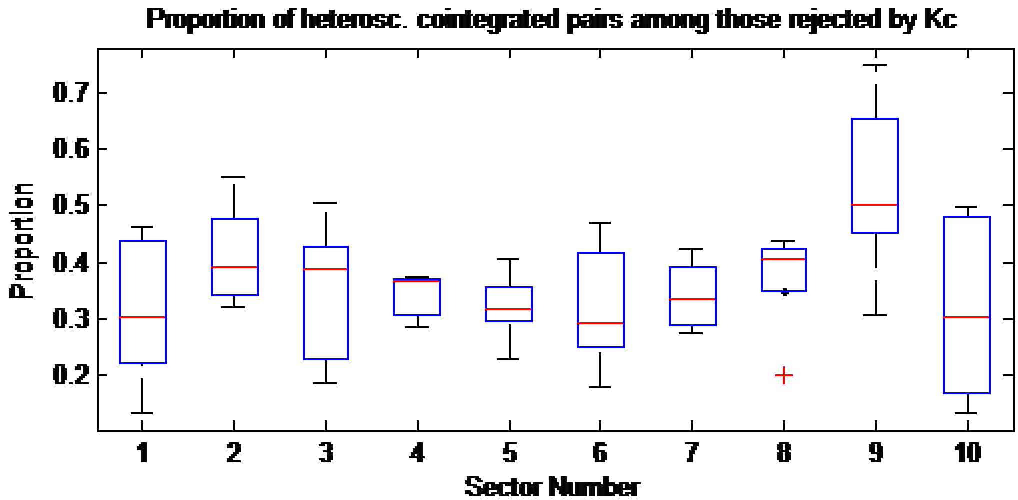


by EG is then calculated as a proportion of the pairs that are detected by S_{nc}.

Fig.5 depicts the proportions of pairs that EG rejects out of the pairs that SC detects. The figure shows that in Sector 2 for example, as high as 50% of the cointegrating pairs that are detected by SC are actually rejected by EG at one time section, while 'only' about 28% are rejected in other time section. With the exception of Sector 9, the medians of all sectors are above 30%, indicating that the EG methodology fails to detect a large proportion of cointegrating pairs.

Next, still focusing on those cointegrating pairs that are detected by S_{nc} but rejected by K_c, we show empirically that a potential reason that K_c fails to detect a larger number of pairs is due to the fact that many of these pairs are heteroscedastically cointegrated and not cointegrated in a stationary fashion, as mandated by the EG framework. As noted in Section 2 the statistic S_{hc} is designed to distinguish between stationary and heteroscedastic cointegration.

Figure 5.6:

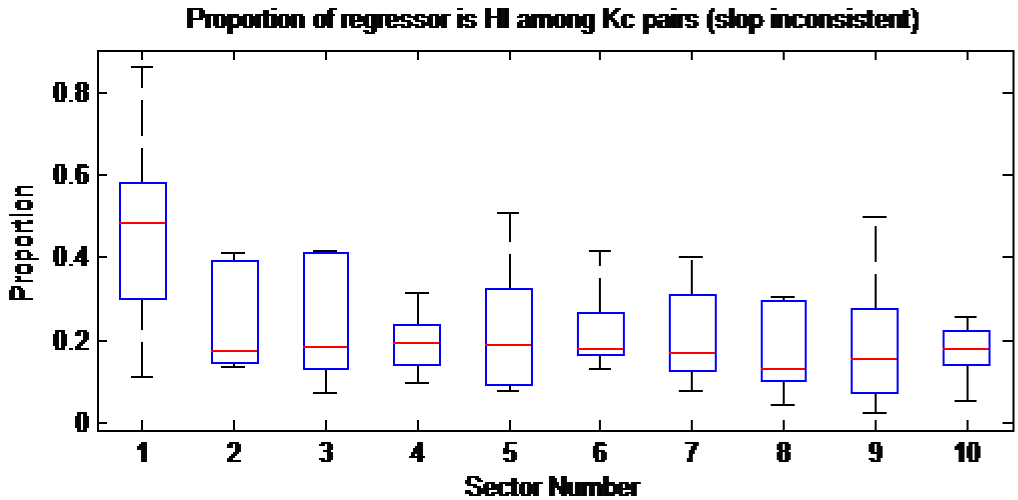


The sector medians in Fig.6 show that over 30% of the pairs that detected by S_{nc} but rejected by K_c are in fact found to be heteroscedastically cointegrated. This highlights the fact that the conventional method largely rejects cointegration in the presence of heteroscedastic integration.

Last, in order to yet further highlight the potential drawback of using the EG, the focus is turned this time to those cointegrating pairs that are detected by both K_c and S_{nc}. As detailed in Section 5.2.2, when the regressor is heteroscedastically integrated, the long-run parameters estimated by (EG) OLS are inconsistent. In each sector at each time section, we look at the proportion of the pairs where the regressor is found to be heteroscedastically integrated according to the S_{hi} test.

The medians proportions in Fig.7 show that about 20% of the cointegrating pairs that are detected by both K_c and S_{nc} have HI regressors. This indicates that a significant proportion of the long-run parameters that are estimated by

Figure 5.7:



EG method is subject to inconsistencies and therefore may not be reliable.

5.4 Conclusion

The empirical findings above suggest that the framework of stochastic integration detects cointegrating relationships in the US equity market. The nonlinear generalization of the framework detects a significantly larger number of cointegrating pairs. The findings are in line with the simulation results reported in HML. The fact that cointegrating pairs of stocks can be found in the equity market points out that relative pricing is an effective pricing methodology when applied to a specific set of stocks. Given cointegrating relationship, relative pricing implies that the price of one stock can be inferred from the price of the other. Granger (1986) states that assets in an efficient market cannot be cointegrated; if they were, there would be a market inefficiency since there would be Granger causality running at least in one direction and thus one price could be used to predict the other. Invoking Granger's definition, our analysis shows that markets are not efficient. Stochastic cointegration is able to detect this type of inefficiency better than the EG method and provides traders with a better way to statistically arbitrage any significant deviations from the long-run relationship. A possible extension of this research would be to simulate a trading strategy that builds on the SC approach to screen for pairs of stocks that exhibit long-run relationship. Then compute the returns from those pairs that significantly deviated from their cointegrating relationship over a historical sample period and compare it to the returns from a buy-and-hold investment strategy. The efficient market hypothesis suggests that the mean returns from each strategy should be equal.

Appendix

Sector 1: Energy and consumption

Sector 2: Material/labor

Sector 3: Industrials

Sector 4: Consumer Discretionary

Sector 5: Consumer Staples

Sector 6: Health Care/liability

Sector 7: Financials

Sector 8: Information Technology

Sector 9: Telecommunication Services

Bibliography

Alexander, C.O., "Correlation and cointegration in energy markets," in "Managing Energy Price Risk," 2nd edition ed., RISK Publications, 1999, pp. 291–304.

Ayat, Leila and Peter Burridge, "Unit root tests in the presence of uncertainty about the non-stochastic trend," *Journal of Econometrics*, 2000, *95* (1), 71 – 96.

Balke, N.S. and T.B. Fomby, "Threshold cointegration," *International Economic Review*, 1997, *38* (3), 627–645. cited By (since 1996) 292.

Becker, Ralf, Walter Enders, and Junsoo Lee, "A Stationarity Test in the Presence of an Unknown Number of Smooth Breaks," *Journal of Time Series Analysis*, 2006, *27* (3), 381–409.

Berk, K.N., "Consistent autoregressive specteral estimates," *The Annals of Statistics*, 1974, *2* (3), 489–502.

Bierens, Herman J., "Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate," *Journal of Econometrics*, November 1997, *81* (1), 29–64.

Bossaerts, P., "Common Nonstationary Components of Asset Prices," *Journal of Economic Dynamics and Control*, 1988, *12*, 347–364.

___ **and R. Green**, "A General Equilibrium Model of Changing Risk Premia: Theory and Evidence," *Review of Financial Studies*, 1989, *2*, 467–493.

Busetti, F. and A.M.R. Taylor, "Testing against stochastic trend in the presence of variance shifts," *Journal of Business and Economic Statistics*, 2003, *21*, 510–531.

Campbell, J.Y. and R.J. Shiller, "Cointegration and tests of present value models," *Journal of Political Economy*, 1987, *95*, 1062–1088.

Cavaliere, G., "Unit root tests under time-varying variances," *Econometric Reviews*, 2004, *23*, 259–292.

Cavaliere, Giuseppe and A.M. Robert Taylor, "Stationarity tests under time-varying second moments," *Econometric Theory*, 2005, *21* (06), 1112–1129.

___ **and** ___ , "Testing for unit roots in time series models with non-stationary volatility," *Journal of Econometrics*, 2007, *140* (2), 919 – 947.

___ **and** ___ , "Bootstrap unit root tests for time series with nonstationary volatilitY," *Econometric Theory*, 2008, *24* (01), 43–71.

___ **, David I. Harvey, Stephen J. Leybourne, and A.M. Robert Taylor**, "Testing for unit roots in the presence of a possible break in trend and nonstationary volatility," *Econometric Theory*, 2011, *FirstView*, 1–35.

Chen, Z. and P. Knez, "Measurement of Market Integration and Arbitrage," *Review of Financial Studies*, 1995, *8*, 287–325.

Choi, In, "Unit root tests for panel data," *Journal of International Money and Finance*, 2001, *20* (2), 249 – 272.

——, "Combination unit root tests for cross-sectionally correlated panels," in Dean Corbae, ed., *Econometric Theory and Practice: Frontiers of analysis and applied research*, Cambridge University Press, 2006, chapter 12.

Daihes, Oron, "Detecting level breaks in the presence of non-stationary volatility," *Working Paper*, 2011.

Dickey, D. A. and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *J. Am. Stat. Assoc.*, 1979, *74*, 427–431.

Elliott, G., T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, 1996, *64*, 813–836.

Engle, R. and C. Granger, "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 1987, *55*, 251–276.

Gatev, E.G., W.N.Goetzmann, and K.G. Rouwenhorst, "Pairs Trading: Performance of a Relative Value Arbitrage Rule," *Review of Financial Studies*, 2006, *19*, 797–827.

Granger, C.W.J., "Developments in the Study of Cointegrated Economic Variables," *Oxford Bulletin of Economics and Statistics*, 1986, *48*, 231–228.

Grilli, Enzo R. and Maw Cheng Yang, "Primary Commodity Prices, Manufactured Goods Prices, and the Terms of Trade of Developing Countries: What the Long Run Shows," *The World Bank Economic Review*, 1988, *2* (1), 1–47.

Hadri, Kaddour, "Testing for stationarity in heterogeneous panel data," *Econometrics Journal*, 2000, *3* (2), 148–161.

Hamilton, James D., *Time Series Analysis*, Princeton, NJ: Princeton University Press, 1994.

Harris, David, Brendan McCabe, and Stephen Leybourne, "Stochastic cointegration: estimation and inference," *Journal of Econometrics*, 2002, *111* (2), 363 – 384.

___ , ___ , and ___ , "Some limit theory for autocovariances whose order depends on sample size," *Econometric Theory*, 2003, *19* (05), 829–864.

___ , David I. Harvey, Stephen J. Leybourne, and A.M. Robert Taylor, "Testing for a unit root in the presence of a possible break in trend," *Econometric Theory*, 2009, *25* (Special Issue 06), 1545–1588.

___ , Stephen Leybourne, and Brendan McCabe, "Panel Stationarity Tests for Purchasing Power Parity With Cross-Sectional Dependence," *Journal of Business and Economic Statistics*, 2005, *23* (4), 395–409.

Harvey, David I., Stephen J. Leybourne, and A.M. Robert Taylor, "Simple, robust, and powerful tests of the breaking trend hypothesis," *Econometric Theory*, 2009, *25* (Special Issue 04), 995–1029.

____, ____, and ____, "Unit Root Testing in Practice: Dealing with Uncertainty over the Trend and Initial Condition," *Econometric Theory*, 2009, *25* (03), 587–636.

____, ____, and ____, "Robust methods for detecting multiple level breaks in autocorrelated time series," *Journal of Econometrics*, 2010, *157* (2), 342 – 358.

____, ____, and ____, "Testing for Unit Roots and the Impact of Quadratic Trends, with an Application to Relative Primary Commodity Prices," *Econometric Reviews*, 2011, *30* (5), 514–547.

Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin, "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 2003, *115* (1), 53 – 74.

J., Jr. Ingersoll, *Theory of Financial Decision-Making*, NJ: Rowman and Littlefield, 1987.

Jegadeesh, N. and S. Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 1993, *48*, 65–91.

Kapetanios, George, Yongcheol Shin, and Andy Snell, "Testing for a unit root in the nonlinear STAR framework," *Journal of Econometrics*, 2003, *112* (2), 359 – 379.

Kim, Tae-Hwan, Stephen Leybourne, and Paul Newbold, "Unit root tests with a break in innovation variance," *Journal of Econometrics*, 2002, *109* (2), 365 – 387.

Kuan, Chung-Ming and Kurt Hornik, "The generalized fluctuation test: A unifying view," *Econometric Reviews*, 1995, *14* (2), 135–161.

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

Leisch, Friedrich, Kurt Hornik, and Chung-Ming Kuan, "Monitoring structural changes with the generalized fluctuation test," *Econometric Theory*, 2000, *16* (06), 835–854.

Maddala, G. S. and Shaowen Wu, "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics*, 1999, *61* (S1), 631–652.

McCabe, Brendan, Stephen Leybourne, and David Harris, "A residual-based test for stochastic cointegration," *Econometric Theory*, 2006, *22* (03), 429–456.

O'Connell, Paul G. J., "The overvaluation of purchasing power parity," *Journal of International Economics*, 1998, *44* (1), 1 – 19.

Ouliaris, Sam, Peter C.B. Phillips, and Joon Y. Park, "Testing for a Unit Root in the Presence of a Maintained Trend," in B. Ray, ed., *Advances in Econometrics and Modelling*, Kluwer, Dordrecht, 1989, pp. 6–28.

Perron, P., "The Great Crash, the Oil Price Shock and the unit root hypothesis," *Econometrica*, 1989, *57*, 1361–1401.

——, "Dealing with structural breaks," in K. Patterson and T.C. Mills, eds., *Econometric Theory*, Vol. 1 of *Palgrave Handbook of Econometrics*, Palgrave Macmillan, 2006, pp. 278–352.

—— **and T. Vogelsang**, "Testing for a unit root in a time series with a changing mean: corrections and extensions," *Journal of Business and Economic Statistics*, 1992, *10*, 467–470.

Perron, Pierre and Xiaokang Zhu, "Structural breaks with deterministic and stochastic trends," *Journal of Econometrics*, 2005, *129* (1-2), 65 – 119. Modelling structural breaks.

—— **and Zhongjun Qu**, "A simple modification to improve the finite sample properties of Ng and Perron's unit root tests," *Economics Letters*, 2007, *94* (1), 12 – 19.

Pfaffenzeller, Stephan, Paul Newbold, and Anthony Rayner, "A Short Note on Updating the Grilli and Yang Commodity Price Index," *The World Bank Economic Review*, 2007, *21* (1), 151–163.

Phillips, Peter C. B., "New Tools for Understanding Spurious Regressions," *Econometrica*, November 1998, *66* (6), 1299–1326.

Pippenger, Michael K. and Gregory E. Goering, "PRACTITIONERS CORNER: A Note on the Empirical Power of Unit Root Tests under Threshold Processes," *Oxford Bulletin of Economics and Statistics*, 1993, *55* (4), 473–481.

Prebisch, Ral, *The economic development of Latin America and its principal problems*, Vol. 7, United Nations Dept. of Economic Affairs, 1950.

Ross, S, "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 1976, *13*, 341–360.

——, "Risk, return and arbitrage," in I. Friend and J. Bicksler, eds., *Risk Return in Finance*, Cambridge, MA: Ballinger, 1976, pp. 341–360.

Saikkonen, P., "Asymptotically efficient estimation of cointegration regressions," *Econometric Theory*, 1991, *7*, 1–21.

Sensier, Marianne and Dick van Dijk, "Testing for Volatility Changes in U.S. Macroeconomic Time Series," *Review of Economics and Statistics*, 2004, *86* (3), 833–839.

Shin, Y., "A residual based test of the null of cointegration against the alternative of no cointegration," *Econometric Theory*, 1994, *10*, 91–115.

Singer, H., "Comments to the terms of trade and economic development," *Review of Economics and Statistics*, 1950, *40*, 84–89.

Sollis, R, *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, Emerald, 2008.

Stock, James H., "Unit roots, structural breaks and trends," in Robert F. Engle and Daniel L. McFadden, eds., *Handbook of Econometrics*, Vol. 4, Elsevier, 1994, pp. 2739–2840.

____ **and Mark W Watson**, "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business and Economic Statistics*, 1996, *14* (1), 11–30.

____ **and Mark W. Watson**, "A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series," in Robert F. Engle and Hal White, eds., *Cointegration, Causality and Forecasting: a Festschrift in Honour of Clive W.J. Granger*, Vol. 1, Oxford University Press, 1999, pp. 1–44.

Xiao, Zhijie, "Quantile Cointegrating Regression," *Journal of Econometrics*, 2009, *150*, 248–260.