Economic Issues, Vol.8, Part 1, 2003

Asymmetric Unit Root Tests in the Presence of Innovation Variance Breaks: Threshold versus Consistent-Threshold Estimation

Steven Cook and Neil Manning[1]

ABSTRACT

Kim et al. *(2002) demonstrate that the Dickey-Fuller unit root test can experience severe size distortion when a large decrease in the innovation variance occurs early in the sample period, leading to spurious rejection of the null. We extend this analysis to the case of spurious identification of asymmetric stationarity by the MTAR test of Enders and Granger (1998) under similar circumstances. In terms of unit root testing, the properties of the MTAR test are inferior to those of the Dickey-Fuller test. However, the MTAR test with consistent-threshold estimation outperforms both the Dickey-Fuller and the original MTAR tests when considering the unit root hypothesis; size distortion being dramatically reduced. The consistent MTAR test is also to be preferred to the original MTAR test when testing the joint hypothesis of non-stationarity and symmetry since the original test can display considerable undersizing. However, the size of the consistent MTAR test is approximately nominal in all experiments except when extreme changes in innovation variance occur towards the beginning of the sample period.*

1. INTRODUCTION

THE SEMINAL STUDY of Perron (1989) has prompted a large literature addressing the issue of testing for unit roots in the presence of structural breaks (see, *inter alia*, Bai *et al* 1998; Bai and Perron 1998; Banerjee *et al* 1992; Perron 1989, 1990; Zivot and Andrews 1992). Consequently, it has long been known that the Dickey-Fuller (DF) (1979) test can fail to reject the null of non-stationarity when a series is in fact stationary but subject to a structural break. However, more recent research has considered the converse phenomenon of the spurious inference of stationarity when a break occurs under the null. The results of Leybourne *et al* (1998) show that when an I(1) series experiences a break in either its level or drift early in the sample period, the DF test can experience severe size distortion, leading to the false inference of stationarity.

A further development of the 'converse Perron phenomenon' is provided by Kim *et al* (2002) where the impact of a break in the innovation variance of a time series upon the properties of the DF test is considered. In contrast to breaks in level and trend, breaks in the innovation variance of a time series are rarely considered in the econometrics literature, particularly for integrated processes.[2] This is perhaps surprising as many financial time-series experience jumps in the innovation variance. Aggarwal *et al* (1999) use the Iterative Cumulative Sums of Squares algorithm of Inclan and Tiao (1994) to demonstrate that large changes in the innovation variance occur in emerging stock markets; similar changes are evident for US dollar bilateral exchange rates for some South East Asian economies, see Manning (2002) for details.

The Monte Carlo results presented by Kim *et al* (2002) show the DF test suffers severe size distortion when an integrated process experiences a break decreasing its innovation variance early in the sample period, thereby resulting in the spurious rejection of the unit root null hypothesis. In this paper the impact of variance breaks upon more recently proposed asymmetric unit root tests will be considered. Following the results of Enders and Granger (1998), Enders (2001) and Cook and Manning (2002a), attention will be focussed upon the momentum threshold autoregressive (MTAR) asymmetric unit root test under both threshold and consistent-threshold estimation. Consideration of the MTAR model under consistent-threshold estimation adds to the recent work of Cook (2001) where the properties of the standard (not consistent) MTAR test were examined in the presence of level and drift breaks under the null. This paper will therefore consider two issues. First, the rejection of the unit root hypothesis under the alternative MTAR and DF tests will be examined in the presence of innovation variance breaks. Second, the overall performance of the MTAR tests will be analysed by considering the joint rejection of the unit root and symmetry hypotheses. Analysis of the joint hypothesis allows us to consider the extent to which the tests indicate 'spurious asymmetric stationarity'.

This paper will proceed as follows. In section 2 the asymmetric unit root tests to be considered are outlined. In section 3 the Monte Carlo design and results are presented. To illustrate the relevance of the Monte Carlo results, section 4 analyses the time series properties of the Indonesian Rupiah-Sterling spot exchange rate over the period 1998-2001. This series was chosen due to its lack of any significant deterministic trend and its substantial decrease in variance in the first quarter of the sample period. Section 5 provides some concluding remarks.

2. Asymmetric unit root tests

Following Pippenger and Goering (1993), the power of the DF test is known to be low in the presence of asymmetric adjustment. A natural extension is provided by Enders and Granger (1998), in which the DF test is modified to allow the unit root hypothesis to be tested against an alternative of stationarity with asymmetric adjustment.[3] Considering the DF test in its simplest form:

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

the sufficient condition for stationarity[4] is given as $-2 < \rho < 0$. To allow for the possibility of asymmetric adjustment about a stationary attractor, Enders and Granger draw upon the threshold autoregressive approach of Tong (1990) using a Heaviside indicator function to partition y_{t-1}. Two specifications of the Heaviside indicator function are proposed based upon $\{y_t\}$ and $\{\Delta y_t\}$ and these respectively lead to the threshold autoregressive (TAR) and the momentum threshold autoregressive (MTAR) asymmetric unit root tests. As the former has been found to possess little power, only the MTAR test will be examined here, resulting in the following generalisation of (1):

$$\Delta y_t = I_t \rho_1 y_{t-1} + (1 - I_t)\rho_2 y_{t-1} + \xi \tag{2}$$

where I_t is the zero-one Heaviside indicator function:

$$I_t = \begin{cases} 1 \; if \; \Delta y_{t-1} \geq 0 \\ 0 \; if \; \Delta y_{t-1} < 0 \end{cases} \tag{3}$$

Asymmetric adjustment about a stationary attractor is then present if $\rho_1 \neq \rho_2$ given prior rejection of the unit root. In the above example, the implicit attractor is $y_t = 0$. However, in some cases it may be more appropriate to re-define the attractor as some other constant or as a trend, in which case the series $\{y_t\}$ is regressed upon the relevant deterministic terms to derive a transformed series $\{\tilde{y}_t\}$ to which (2) and (3) are then applied.

To test the unit root hypothesis, the null H_0: $\rho_1 = \rho_2 = 0$ is examined, values of the test statistics being compared with specifically calculated critical values. Should the null be rejected, the distribution of the adjustment coefficients ρ_1 and ρ_2 converges to a multivariate Normal distribution according to Tong (1983, 1990). Consequently, under the maintained assumption of stationarity, symmetry $\{\rho_1 = \rho_2\}$ can be tested using a conventional F-statistic.

Recently Enders (2001) and Cook and Manning (2002a) have reconsidered the MTAR test following regression upon a constant, denoted as Φ_μ, noting that the two-step procedure effectively demeans the series. The version of (2) estimated is therefore:

$$\Delta y_t = I_t \rho_1 (y_{t-1} - \bar{y}) + (1 - I_t) \rho_2 (y_{t-1} - \bar{y}) + \xi \tag{4}$$

with the indicator function as defined in (3). In the presence of asymmetric adjustment the mean will be a biased estimator of the threshold, as ρ_1 and ρ_2 differ. In such circumstances, the grid search approach of Chan (1993) and Chan and Tsay (1998) should be followed to obtain a super-consistent value of the threshold. Considering MTAR adjustment, consistent-threshold estimation requires reordering the $\{\Delta y_t\}$ process in terms of increasing size, i.e. $\Delta y_2^0 < \Delta y_2^0 < \ldots < \Delta y_T^0$, with a subset of the Δy_j^0 being considered in turn as a potential threshold. For each $\Delta y_j^0 = \tau$, the following equation is estimated:

$$\Delta y_t = I_t \rho_1 (y_{t-1} - \tau) + (1 - I_t) \rho_2 (y_{t-1} - \tau) + \eta_t \tag{5}$$

where:

$$I_t = \begin{cases} 1 \text{ if } \Delta y_{t-1} \geq \tau \\ 0 \text{ if } \Delta y_{t-1} < \tau \end{cases} \tag{6}$$

The value of Δy_j^0 delivering the minimum residual sum of squares ($\Sigma \eta_t^2$) is then selected as the consistent threshold, with the resultant test being denoted as $\Phi^*(c)$. However, after re-ordering the series, a number of the highest and lowest ordered values of observations on $\{\Delta y_t\}$ are excluded from the grid search identifying τ to ensure sufficient degrees of freedom are present in each regime. It should be noted that Enders (2001) and Cook and Manning (2002a) draw differing conclusions concerning the relative powers of the MTAR test under threshold and consistent-threshold estimation. More precisely, the Monte Carlo analysis in Cook and Manning (2002a) reverses the counterintuitive finding in Enders (2001), namely that consistent-threshold estimation results in lower power; intuitively one would expect the use of a consistent-threshold to result in higher power.[5] Cook and Manning's findings are supported by the empirical results in Enders and Siklos (2001) for the MTAR cointegration test, and in Cook and Holly (2002) for the Granger-Lee asymmetric error correction model. In both cases, use of consistent-threshold estimation as opposed to estimation with an imposed threshold allows the respective nulls to be rejected. Crucially, these studies do not conduct simulation analyses examining the power properties of consistent-threshold estimation and only report their empirical test findings.

However in this paper, simulation exercises examine spurious rejection of the unit root hypothesis by the $\Phi_\mu{}^*$, $\Phi^*(c)$ and DF tests in the presence of innovation variance breaks. In addition, the simulations also examine the joint rejection of the unit root and symmetry hypotheses by the $\Phi_\mu{}^*$, $\Phi^*(c)$ tests.

3. Monte Carlo simulation

3.1 Experimental design

To examine the possible size distortion of the $\Phi_\mu{}^*$, $\Phi^*(c)$ and DF tests is the presence of innovation error variance breaks, the following data generation process given by (7)-(10) is employed:

$$y_t = y_{t-1} + \xi_t \qquad t = 1,\dots,T \tag{7}$$

$$\xi_t = \sigma_t \eta_t \tag{8}$$

$$\eta_t \sim i.i.d. \quad \mathrm{N}(0,1) \tag{9}$$

$$\sigma_t^2 = \begin{cases} \sigma_1^2 \text{ for } t \le T_B \\ \sigma_2^2 \text{ for } t > T_B \end{cases} \tag{10}$$

where T_B is the period when the break in variance occurs. The error series $\{\eta_t\}$ is generated using the RNDNS procedure in the Gauss programming language version 3.2.13, with the initial value $\{y_0\}$ set equal to zero. All experiments are performed over 10,000 replications with the first 100 observations of the series discarded to remove the influence of initial conditions.

Kim *et al* (2000) show the size distortion of the DF test depends on the magnitude of the decrease in variance and the period during which it occurs. Denoting the ratio σ_2/σ_1 as δ we employ the range of values $\delta \in \{0.2,0.4,0.6,0.8\}$ setting $\sigma_1 = 1$ without loss of generality. Following the power analyses of Cook and Manning (2002a), Enders and Granger (1998) and Enders (2001), a sample size of $T = 100$ is employed throughout. However, we extend the design in Kim *et al* (2000) by permitting the break in variance to occur at all possible points in the sample, that is $T_B = \{1, 2,\dots, 98, 99\}$. With the data thus generated, the series $\{\Delta y_t\}$ is then reordered in increasing size to allow application of the $\Phi^*(c)$ test. When minimising the sum of squares to optimise the value of the threshold τ, the 15 initial and 15 final values of the reordered Δy_t series are omitted from this grid search procedure to ensure sufficient degrees of freedom exist in each regime. To calculate empirical frequencies at the 5 per cent nominal level for the $\Phi^*(c)$ test, (5) and (6) are employed with the critical value being taken from Cook and

Manning (2002a).

When considering the $\Phi_\mu{}^*$ test, the series $\{y_0\}$ is demeaned by prior regression on a constant before applying (2) and (3) to the resultant revised series. To calculate empirical rejection frequencies, the critical value is taken from Enders (2001). When examining the joint rejection of the unit root and symmetry hypotheses using the $\Phi_\mu{}^*$ and $\Phi^*(c)$ tests, we employ a standard F-distribution in testing the secondary symmetry hypothesis, as discussed above.

For comparative purposes, the standard τ test is calculated by running the following regression:

$$\Delta y_t = \mu + \rho y_{t-1} + \varepsilon_t \tag{11}$$

with the appropriate critical value being taken from Fuller (1976).

3.2 Results

Given the large range of breakpoints considered, we follow the approach of Davidson and MacKinnon (1996) and present the results of the Monte Carlo analysis graphically to ease interpretation and avoid a plethora of repetitive tabulations. Figures 1(a)-1(d) summarise the performance of the three tests across all possible breakpoints in the sample period for the four alternative ratios (δ) when considering the null of a unit root. From these figures, it is apparent that for both the τ_μ and $\Phi_\mu{}^*$ tests severe size distortion can occur when large decreases in innovation variance are experienced in the initial part of the sample period.[6] As expected, when δ is increased from 0.2 to 0.8, size distortion decreases dramatically. However, it is also apparent that $\Phi_\mu{}^*$ suffers greater size distortion than does the τ_μ test. In fact, size distortion of the $\Phi_\mu{}^*$ test reaches a maximum when $\{\delta, T_B\}$ with a value of 57.93 per cent at the 5 per cent nominal level. Now consider the performance of the $\Phi^*(c)$ test. The degree of size distortion is dramatically reduced in comparison with the results for the $\Phi_\mu{}^*$ and τ_μ tests, spurious rejections occurring only for the smallest values of δ. Maximum distortion now occurs when $\{\delta, T_B\} = \{0.2, 17\}$, the calculated size of 15.99 per cent for the $\Phi^*(c)$ test comparing very favourably to the 56.02 per cent for the $\Phi_\mu{}^*$ test when $\{\delta, T_B\} = \{0.2, 17\}$.

According to the results in Figures 2(a)-2(d) for the $\Phi_\mu{}^*$ and $\Phi^*(c)$ tests, the joint tests also exhibit oversizing when large variance breaks occur early in the sample, as illustrated in Figure 2(a). However, the $\Phi_\mu{}^*$ exhibits undersizing when the same break in variance occurs towards the end of the sample period. Crucially, the performance of the $\Phi^*(c)$ test is more satisfactory than the $\Phi_\mu{}^*$ test when $\delta = 0.2$ in two key respects. First, the probability of spurious asymmetry being inferred is significantly lower when the variance break occurs towards the beginning of the sample and secondly, in contrast to the results for the $\Phi_\mu{}^*$ test, size remains approximately nominal for the $\Phi^*(c)$ test when the variance break occurs towards the end of the sample period. Figures 2(b) and 2(c) illustrate the performance of the two tests for intermediate values of δ and two characteristics are immediately apparent. When the variance break occurs early in the sample period, application of the $\Phi^*(c)$ test is slightly more likely to result in spurious rejections of the null. However, when the same variance break occurs towards the end of the sample the $\Phi^*(c)$ test once more has nominal size, yet the $\Phi_\mu{}^*$ test exhibits considerable undersizing. This undersizing of the $\Phi_\mu{}^*$ test is clearly evident in Figures 2(c) and 2(d) irrespective of when the break in variance occurs and is also apparent when there is no break in variance whatsoever.[7] In contrast, Figure 2(d) shows the $\Phi^{*}(c)$ test to have approximately nominal size and this satisfactory performance also occurs when the innovation variance is homoscedastic.

Figure 1a: unit root testing, δ=0.2

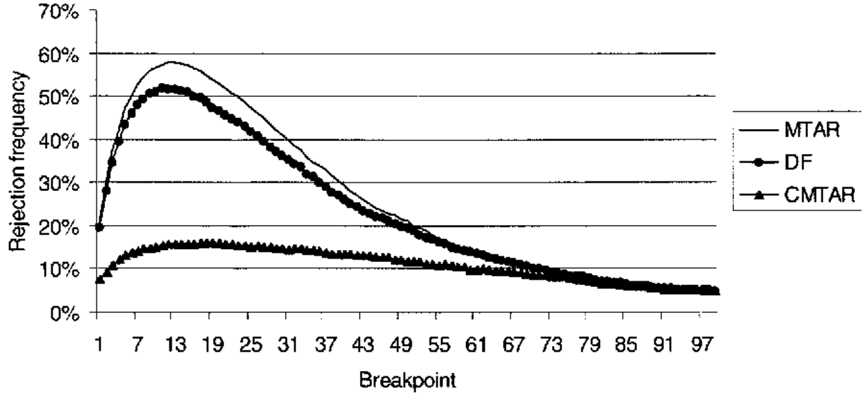


Figure 1b: unit root testing, δ=0.4

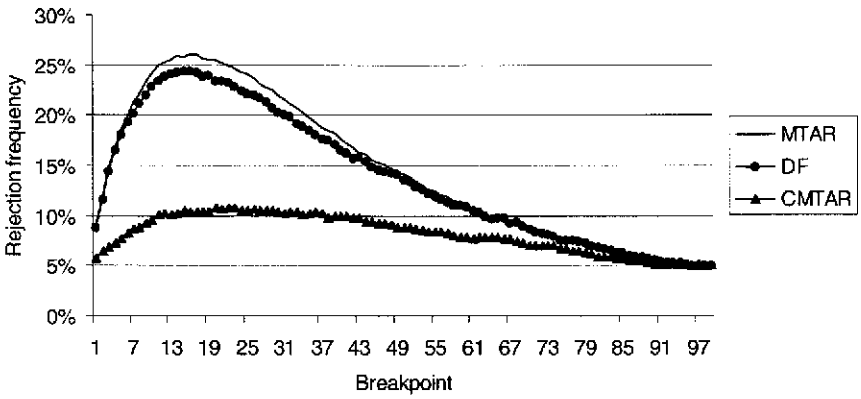


Figure 1c: unit root testing, δ=0.6

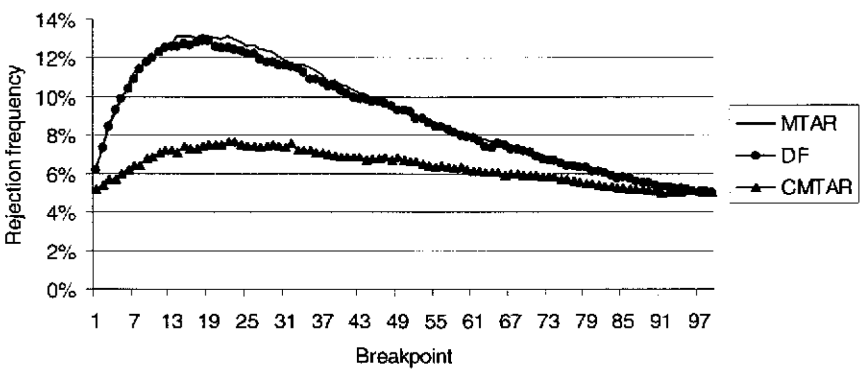


Figure 1d: unit root testing, δ=0.8

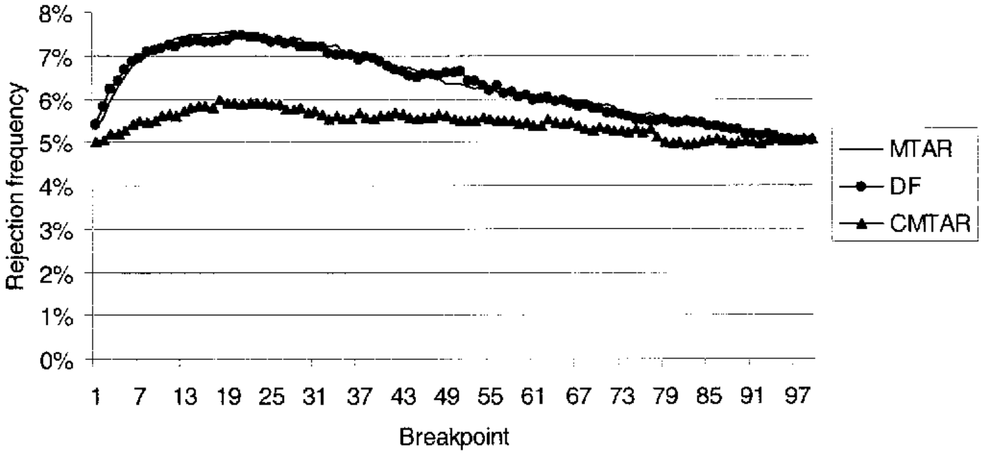


Figure 2a: joint hypothesis testing, δ=0.2

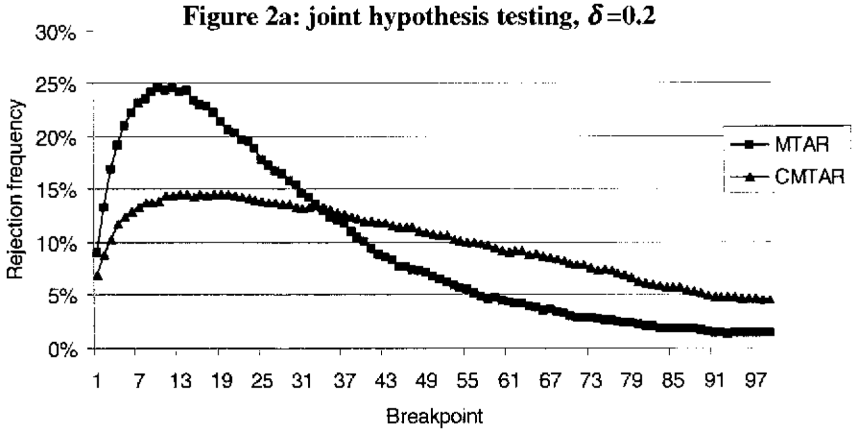


Figure 2b: joint hypothesis testing, δ=0.4

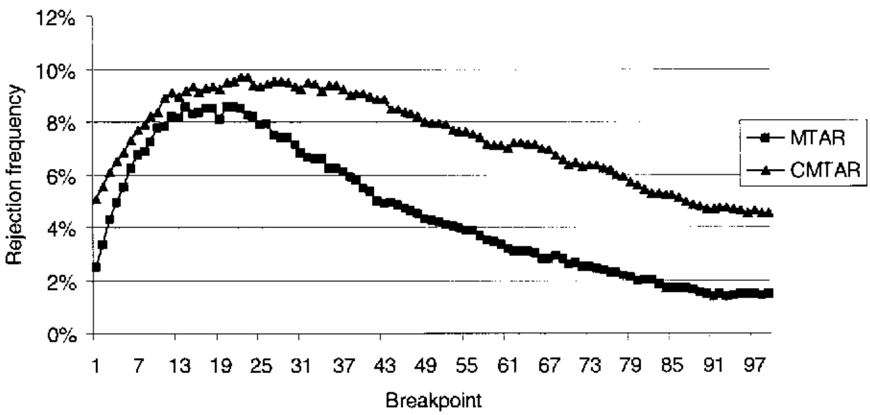


Figure 2c: joint hypothesis testing, δ =0.6

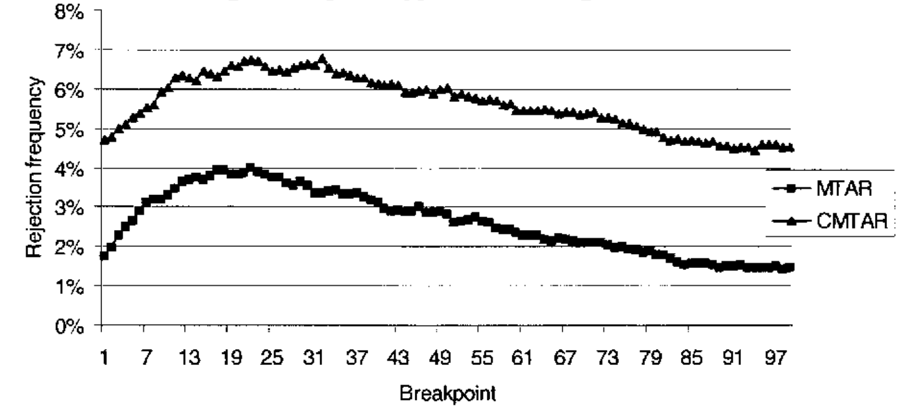
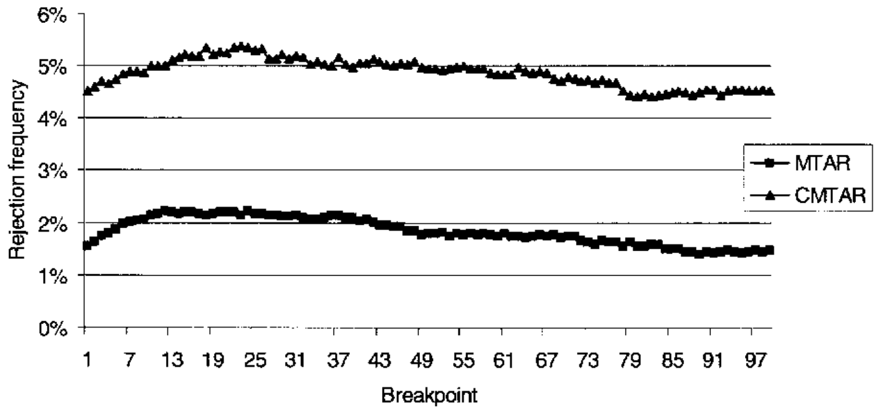


Figure 2d: joint hypothesis testing, δ =0.8

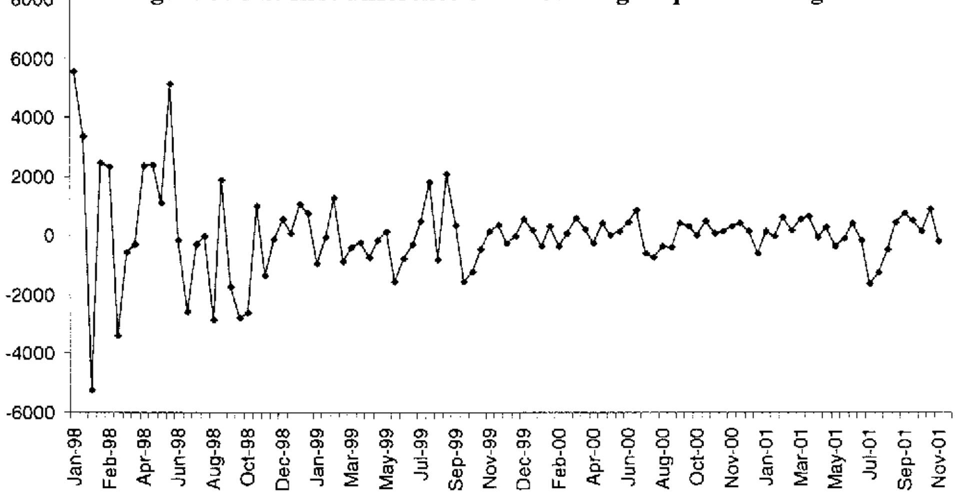


4. EMPIRICAL ILLUSTRATION

As an illustration of the relevance of the above Monte Carlo results, consider the following analysis of the spot Rupiah-Sterling exchange rate, the data being sampled fortnightly over the period January 1998 to November 2001 inclusive.[8] This series has characteristics which conform to the above Monte Carlo design, namely, it exhibits no significant trend over the sample and there is a substantial decrease in variance between the sub-periods 1998 and 1999-2001. In terms of the above Monte Carlo parameters, the sample has the following characteristics: δ = 0.36, T = 100 and T_B = 25. These particular values relate to the more extreme cases in the above Monte Carlo simulations, and illustrate the relevance of our design to empirical applications. Given these values of T, T_B and δ, the Monte Carlo results suggest that severe distortion occurs in the Dickey-Fuller and original MTAR tests, but this is not the case for the consistent form of the MTAR test, namely $\Phi^*(c)$.

Figure 3 presents the exchange rate data in difference form, and τ_μ, Φ_μ^* and $\Phi^*(c)$ are calculated as above in Section 2. The standard tests both reject the null of a unit root: τ_μ = -3.08, with 5 per cent critical value = -2.89 and $\Phi_{\mu}{}^{*}$ = 5.19, the relevant 5 per cent critical value being 5.02. One may be tempted to infer therefore that the Rupiah spot rate possesses asymmetric stationarity given the above value of the $\Phi_{\mu}{}^{*}$ test. However, the $\Phi_{\mu}{}^{*}$ test of symmetry, which is less prone to distortion in the presence of large variance breaks than the equivalent unit root test, fails to indicate any asymmetric reaction whatsoever since $F(1,98)$ = 0.80. In contrast to the above results, the consistent MTAR test fails to reject the null of a unit root: $\Phi_{\mu}{}^{*}$ = 0.97, a test value considerably below the 5 per cent critical value of 5.23.

Figure 3: The first difference of the Sterling-Rupiah exchange rate



This example therefore illustrates the relevance of our Monte Carlo experiments; the time series properties of actual data may be incorrectly classified using standard tests when variance breaks are present.

5. Conclusion

This paper has extended previous results concerning the impact of changes in the innovation variance on tests of the unit root hypothesis to specifically consider asymmetric unit root tests and their properties. We find that in common with the standard Dickey-Fuller unit root tests, asymmetric unit root tests can also suffer severe size distortion if large decreases in the innovation variance occur early in the sample period. This is particularly the case if the threshold is imposed; if the threshold is left unconstrained and consistent-threshold estimation is employed the properties of the resultant test are notably improved. In fact, we find the $\Phi_{\mu}{}^{*}$ is even outperformed by the standard τ_{μ} test when considering rejection of the null of a unit root.

We also consider the secondary hypothesis of symmetry and once again the resultsdiffer between threshold and consistent-threshold estimation. We find that spurious inference of asymmetric stationarity is also likely if a large decrease in the innovation variance occurs early in sample. However, the $\Phi_{\mu}{}^{*}$ joint test of stationarity and symmetry can be severely undersized in other circumstances, most commonly when the break in the innovation variance occurs in the second half of the sample period or when the break is small or absent. Of the tests considered, we find that $\Phi^*(c)$ possesses the best properties both as unit root test and as a test of the joint hypothesis of a unit root and symmetry. These results, taken in conjunction with power calculations reported in Cook and Manning (2002a), suggest the MTAR test in its consistent form be employed by practitioners when undertaking this type of hypothesis testing. These recommendations are also supported by the results of our empirical example, the unit root hypothesis being rejected by both the τ_μ and Φ_μ^* tests but not by our preferred $\Phi^*(c)$ test.

ENDNOTES

1. Department of Economics, University of Wales Swansea, Singleton Park, Swansea, SA2 8PP. Tel: (01792) 295168. E-mail: s.cook@swan.ac.uk, d.n.manning@swan.ac.uk.

2. The papers of Wichern, Miller and Hsu (1976), Hsu (1977), Inclan (1993) and Inclan and Tiao (1994) can be considered as the rare exceptions to this where a break in innovation variance is considered in more general circumstances. Kim *et al* (2002) cite Hamori and Tokihisa (1997) as the only case where the impact of a variance break upon an integrated series is examined.

3. Such an alternative hypothesis has a natural appeal given the growing theoretical literature suggesting asymmetric adjustment in a range of economic variables see, *inter alia*, Ball and Mankiw (1994), Gale (1996), Krane (1994).

4. In the fractional integration literature, series are non-stationary for all fractional differences $\delta > 0.5$. However, in the present context, we consider series to be stationary if $\delta < 1$.

5. This is not the case if the true threshold value equals zero, in which case the MTAR test, which does not involve estimation of the threshold, would have the higher power.

6. This echoes results in Kim *et al* (2002).

7. See Cook and Manning (2002b).

8. Source: Datastream. Series code: INDORUP.

REFERENCES

Cook S (2001) 'Asymmetric unit root tests in the presence of structural breaks under the null', *Economics Bulletin*, 3 (6), 1-10.

Cook S and Holly S (2002) 'Threshold specification for asymmetric error correction models', *Applied Economics Letters*, 9, 711-13.

Cook S and Manning N (2002a) 'The power of asymmetric unit root tests under threshold and consistent-threshold estimation', mimeo, University of Wales, Swansea.

Cook S and Manning N (2002b) 'The relative sizes of alternative asymmetric unit root tests', mimeo, University of Wales, Swansea.

Davidson R and MacKinnon J (1998) 'Graphical methods for investigating the size and power of hypothesis tests', *Manchester School*, 66, 1-26.

Dickey D and Fuller W (1979) 'Distribution of the estimators for autoregressive time series with a unit roo', *Journal of the American Statistical Association*, 74, 427-431.

Dixit A (1992) 'Investment and hysteresis', *Journal of Economic Perspectives*, 6, 107-132.

Enders W (2001) 'Improved critical values for the Enders-Granger unit root test', *Applied Economics Letters*, 8, 257-261.

Enders W and Granger C (1998) 'Unit root tests and asymmetric adjustment with an example using the term structure of interest rates', *Journal of Business and Economic Statistics*, 16, 304-311.

Enders W and Siklos P (2001) 'Cointegration and threshold adjustment', *Journal of Business and Economics Statistics*, 19, 166-176.

Fuller W (1976) *Introduction to Statistical Time Series*, New York: Wiley.

Gale D (1996) 'Delay and cycles', *Review of Economic Studies*, 63, 169-98.

Hamori S and Tokihisa A (1997) 'Testing for a unit root in the presence of a variance shift', *Economics Letters*, 57, 245-253.

Hsu S (1977) 'Tests for variance shift at an unknown time point', *Applied Statistics*, 26, 279-284.

Inclan C (1993) 'Detection of multiple changes of variance using posterior odds', *Journal of Business and Economic Statistics*, 11, 289-300.

Inclan C and Tiao G (1994) 'Use of cumulative sums of squares for retrospective detection of changes of variance', *Journal of the American Statistical Association*, 89, 913-923.

Kim T-H, Leybourne S and Newbold P (2000) 'Unit root tests with a break in innovation variance', *Journal of Econometrics*, 109, 365-87.

Krane S (1994) 'The distinction between inventory holding and stock-out costs: implications for target inventories, asymmetric adjustment and the effect of aggregation on production smoothing', *International Economic Review*, 35, 117-136.

Leybourne S, Mills T and Newbold P (1998) 'Spurious rejections by Dickey-Fuller tests in the presence of a break under the null', *Journal of Econometrics*, 87, 191-203.

Maddala G S and Kim I-M (1998) *Unit roots, cointegration and structural change*, Cambridge: Cambridge U P.

Manning N (2002) News and volatility spillovers, asymmetry and non-synchronous trading: interactions among equity markets across four continents', mimeo, University of Wales Swansea.

Perron P (1989) 'The Great Crash, the oil price shock and theunit root hypothesis', *Econometrica*, 57, 1361-1401.

Perron P (1990) 'Testing for a unit root in time series with a changing mean', *Journal of Business and Economic Statistics*, 8, 153-162.

Pippenger M and Goering G (1993) 'A note on the empirical power of unit root tests under threshold processes', *Oxford Bulletin of Economics and Statistics*, 55, 473-481.

Tong H (1983) *Threshold models in non-linear time-series analysis*, New York: Springer-Verlag.

Tong H (1990) *Non-linear time-series: a dynamical approach*, Oxford: Oxford U P.

Wichern D, Miller R and Hsu D (1976) 'Changes of variance in first order autoregressive time series models with an application', *Applied Statistics*, 25, 248-256.

Zivot E and Andrews D (1992) 'Further evidence on the Great Crash, the oil price shock and the unit root hypothesis', *Journal of Business and Economic Statistics*, 10, 251-270.