Economics Bulletin

# Asymmetric unit root tests in the presence of structural breaks under the null

Steven Cook
*University of Wales Swansea*

## *Abstract*

Using Monte Carlo methods, the behaviour of the momentum threshold autoregressive (MTAR) unit root test of Enders and Granger (1998) is examined in the presence of structural breaks under the null. It is found that for level breaks the MTAR test exhibits similar behaviour to that derived by Leybourne et al. (1998) for the Dickey−Fuller (1979) test, with size distortion apparent for early breaks only. In contrast, the results for breaks in drift show the MTAR test to experience severe size distortion when breaks occur both early and late in the sample period. The divergence in results for the MTAR and DF tests is further examined, showing that in the presence of late breaks the MTAR test can lead a practitioner to draw false inferences of both stationarity and asymmetry.

**Citation:** Cook, Steven, (2001) "Asymmetric unit root tests in the presence of structural breaks under the null." *Economics Bulletin*, Vol. 3, No. 6 pp. 1−10
**Submitted:** July 9, 2001. **Accepted:** July 17, 2001.
**URL:** http://www.economicsbulletin.com/2001/volume3/EB−01C10002A.pdf

# 1 Introduction

The use of the Dickey-Fuller (1979,1981) (DF) test to examine the order of integration of economic time series has become a standard feature of applied econometric research. Following the work of Perron (1989) it has long been recognised that the DF test can have very low power when applied to series which are stationary but experience a structural break. In response to the possibility of I(0) series appearing to be I(1) in the presence of structural breaks, a huge literature has subsequently evolved addressing the issue of unit root testing in such circumstances (see, *inter alia*, Bai *et al.* 1998; Bai and Perron 1998; Banerjee *et al.* 1992; Perron 1989, 1990; Zivot and Andrews 1992). However, in contrast to this, Leybourne *et al.* (1998) have recently considered the converse issue of spurious rejections by the DF test when there is a break under the null. The results obtained show that when there is a break in either the level or the drift of a unit root process early in the sample period, the DF test can experience severe size distortion, leading to the false conclusion that an I(1) series is I(0).

In this paper the impact of structural breaks upon the more recently proposed momentum-threshold autoregressive (MTAR) asymmetric unit root test of Enders-Granger (1998) (EG) is considered. The results obtained show that breaks in the *level* of a unit root process have a similar impact on both DF and MTAR unit root tests. However, for breaks in *drift* the MTAR test experiences size distortion for both early and late breaks. It is shown that this can lead to false inferences of stationarity and asymmetry being drawn.

This paper will proceed as follows. In section [2] the EG unit root tests are outlined. Following a discussion of the seminal work of EG, a recent re-evaluation by Enders (2001) is presented. Consideration of this more recent research leads to the present analysis focussing upon just one of the two tests developed by EG. In section [3] Monte Carlo simulation results for the level break case are presented, with section [4] containing results for a break in drift. Section [5] examines the differing results obtained for MTAR and DF tests, with section [6] concluding.

# 2 Asymmetric unit root tests

To extend the familiar DF statistic to allow the unit root hypothesis to be tested against an alternative of stationarity with asymmetric adjustment, EG employ the threshold autoregressive models of Tong (1983, 1990). This allows the lagged levels term of the DF test to be partitioned into positive and negative components. Considering the simplest case excluding deterministic and lagged difference terms, the EG asymmetric testing equation can be given as:

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

where $I_t$ is the zero-one Heaviside indicator function. The use of two specifications for $I_t$ leads to the derivation of two forms of asymmetric unit root test. Under the first approach partitioning is based upon the sign of $y_{t-1}$, while under the second approach partitioning is based upon the sign of the change in $y_{t-1}$. This results in the specifications for $I_t$ given below:

$$I_t = \begin{cases} 1 & \text{if } y_{t-1} \geq 0 \\ 0 & \text{if } y_{t-1} < 0 \end{cases} \tag{2}$$

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

Asymmetric unit root tests combining (1) and (2) are referred to as threshold autoregressive (TAR), while use of (1) and (3) leads to the momentum threshold autoregressive (MTAR) test. To distinguish between these approaches, the TAR and MTAR tests are denoted as $\Phi$ and $\Phi^*$ respectively. Under both methods the unit root hypothesis is tested via the joint significance of $\{\rho_1, \rho_2\}$. While the results of Tong (1983) show $\{\rho_1, \rho_2\}$ to converge to a multivariate Normal distribution under the assumption of stationarity, under the null of a unit root $\{\rho_1, \rho_2\}$ have a non-standard distribution. The unit root hypothesis ($\text{H}_0 : \rho_1 = \rho_2 = 0$) is therefore examined using specifically derived critical values provided by EG. Should the unit root hypothesis be rejected, the process is assumed to be stationary with asymmetric adjustment.

In (1) above, the implicit underlying attractor about which adjustment occurs is zero $(y_t = 0)$. However, in practice it is more appropriate to consider the attractor as either a constant or trend. To employ these attractors, the original series $\{y_t\}$ is regressed upon the relevant deterministic terms. The resulting revised series $\{\widetilde{y}_t\}$ is then employed in the testing equation of (1), giving the following model:

$$\Delta \widetilde{y}_t = I_t \rho_1 \widetilde{y}_{t-1} + (1 - I_t)\, \rho_2 \widetilde{y}_{t-1} + \zeta_t \tag{4}$$

Similarly, the above Heaviside indicator functions are also modified to use the new $\{\widetilde{y}_t\}$ series. In the presence of non-zero attractors the resulting $\Phi$ and $\Phi^*$ employ the subscripts $\mu$ and $T$ to distinguish between the constant and constant & trend cases, respectively.

In recent research Enders (2001) has re-evaluated the seminal work of EG, noting that in the presence of asymmetric adjustment the EG two-step approach will lead to the derivation of a biased estimate of the threshold. Consequently, Enders employs the estimation methods of Chan (1993) and Chan and Tsay (1998) to derive TAR and MTAR unit root tests using 'super-consistent' threshold estimates. A subsequent power analysis of these models by Enders led to two interesting results. First, under TAR asymmetric adjustment, the DF test was found to be more powerful than the TAR test under either consistent or non-consistent threshold estimation. Second, under MTAR adjustment, the original MTAR test was seen to be more powerful than both the consistent MTAR test and the DF test. The results lead to the conclusion that of the rival asymmetric unit root tests, only the MTAR test using non-consistent threshold estimates improves upon the DF test. For this reason, the original MTAR tests ($\Phi_{\mu}^{*}$ and $\Phi_{T}^{*}$) are the only asymmetric unit root tests considered here.[1]

Following the results of Leybourne *et al.* (1998) for the DF test, section [3] examines the behaviour of the $\Phi_{\mu}^{*}$ in the presence of a level break in a unit root process, while section [4] considers the $\Phi_{T}^{*}$ test in the presence of a break in drift.

# 3 Unit root with a break in level

In the presence of a level break in a unit root process, the appropriate test to consider is the $\Phi_{\mu}^{*}$ test.[2] To examine the possible size distortion of the $\Phi_{\mu}^{*}$ test, the following data generation process (DGP) given by (5)-(8) was employed:

$$
\begin{aligned}
y_t &= \alpha s_t(\tau) + \xi_t \qquad t = 1, ...T && (5)\\
\xi_t &= \xi_{t-1} + \eta_t && (6)\\
\eta_t &\sim i.i.d.\ \mathsf{N}(0,1) && (7)\\
s_t(\tau) &= \begin{cases} 0 \text{ for } t \leqslant \tau T \\ 1 \text{ for } t > \tau T \end{cases} \qquad \tau \in (0,1) && (8)
\end{aligned}
$$

The above is therefore the DGP of Leybourne *et al.* (1998). The error series $\{\eta_t\}$ was generated using the RNDNS procedure in the Gauss programming

[1]Enders (2001) also considers the impact of lag order on the critical values of asymmetric unit root tests. However this issue is not relevant in the present context as only non-augmented tests are employed.

[2]Note that in the case of a level break, Leybourne *et al.* (1998) also consider the use of a misspecified DF test which includes a trend term. Results for the use of misspecified models are not presented here.

language. All experiments were performed over 5,000 replications using a sample size of 100 observations ($T = 100$). An additional initial 100 observations were discarded to remove the influence of the initial condition $y_0 = 0$.

To further replicate the experimental design of Leybourne *et al.* (1998) the values $\alpha \in \{2.5, 5, 10\}$ were chosen for the break magnitude. For each replication the $\Phi_\mu^*$ test was estimated with the (false) rejections of the unit root hypothesis noted at the 5% level of significance. Following the arguments of Davidson and MacKinnon (1998), the results of the above experiments are presented graphically in Figure One. From inspection of this graph it can be seen that the behaviour of the $\Phi_\mu^*$ test is similar to that of the DF test. To summarise the results, size distortion is only present for breaks which occur early in the sample period and is greater for larger values of the break magnitude $\alpha$. To illustrate this, the empirical size of the $\Phi_\mu^*$ test for $\alpha = 10$ is 58.96% for $\tau = 0.01$ ($t = 1$), but rapidly approaches nominal size (0.05) as the breakpoint appears later in the sample.

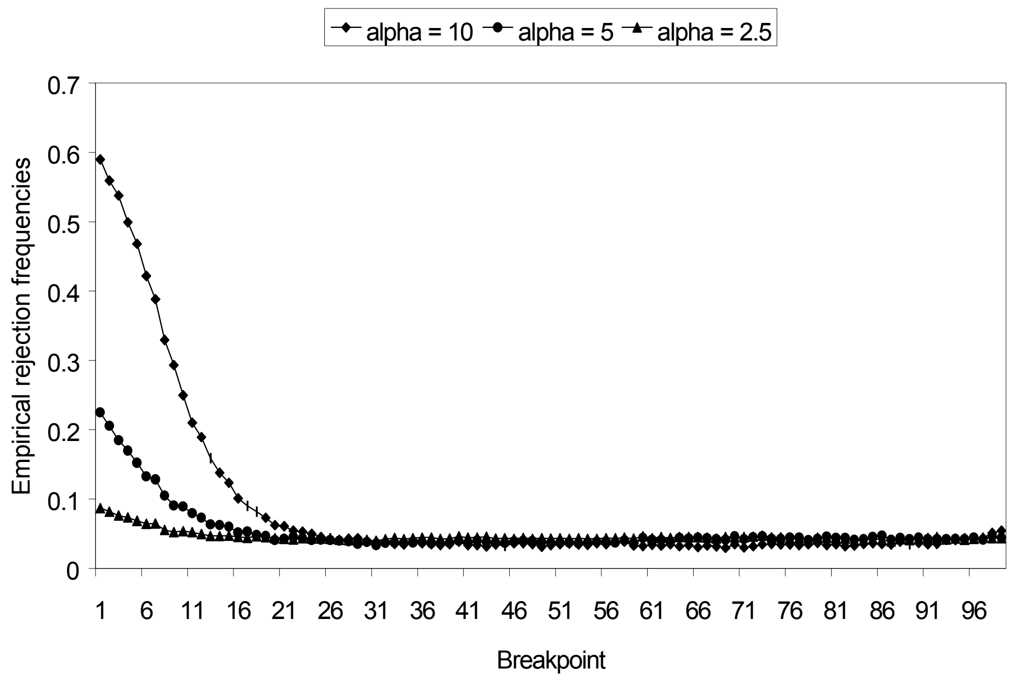


Figure One: MTAR unit root rejections (break in level)

# 4 Unit root with a break in drift

To analyse the break in drift case, the earlier DGP of (5)-(8) is modified as below:[3]

$$
\begin{aligned}
y_t &= \alpha s_t(\tau) + y_{t-1} + \epsilon_t \qquad t = 1, ...T && (9)\\
\epsilon_t &\sim i.i.d.\ \mathsf{N}(0,1) && (10)\\
s_t(\tau) &= \begin{cases} 0 \text{ for } t \leqslant \tau T \\ 1 \text{ for } t > \tau T \end{cases} \qquad \tau \in (0,1) && (11)
\end{aligned}
$$

Following Leybourne *et al.* (1998) the sizes of drift break considered were $\alpha \in \{0.5, 1, 2\}$. The empirical rejection frequencies of the appropriate $\Phi_T^*$ test at the 5% nominal level of significance for each of these cases are presented in Figure Two. From inspection of Figure Two it is immediately apparent that the results for the $\Phi_T^*$ test differ from those for the DF test where breaks in drift are found to cause size distortion only when occurring early in the sample period. For a large break in drift ($\alpha = 2$), size distortion can result when the break occurs either early or late in the sample period. Further considering the case of $\alpha = 2$, the size distortion early in the sample reaches a maximum empirical size of 81.56% for $\tau = 0.14$ ($t = 14$). Although this represents huge distortion it is slightly below the value of 92.9% reported by Leybourne *et al.* (1998) for the DF test. Considering the secondary distortion late in the sample, this peaks at 24.52% for $\tau = 0.85$ ($t = 85$). Given the 5% nominal level employed, this also represents a severe distortion, which interestingly is not present for the DF test.

[3]The treatment of initial conditions, method of random number generation, sample size, and number of replications and discards for the break in drift experiments are the same as for the earlier level break experiments.

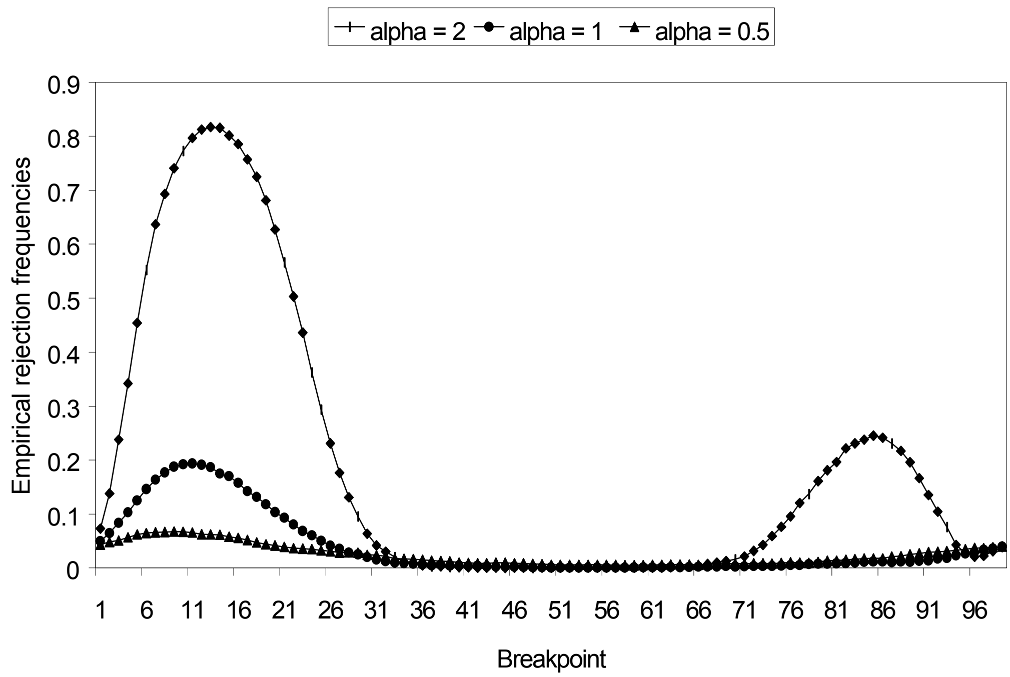


Figure Two: MTAR unit root rejections (break in drift)

# 5 Analysing the differing behaviour of the DF and MTAR tests

The above results show that in the presence of *levels* breaks, the MTAR test exhibits similar behaviour to the DF test with severe size distortion present only when breaks occur early in the sample period. However, for breaks in *drift* the tests behave differently, with the MTAR test experiencing size distortion when breaks occur both early and late in the sample period. To further examine this divergence in behaviour, the case where this is most apparent was reanalysed. Using the break in drift experiment of the previous section with $\alpha = 2$, the MTAR $\Phi_T^*$ and the DF $\tau_\tau$ tests were performed. To analyse the differing behaviour of these tests, the cases where one test rejected the unit root hypothesis but the other did not, were noted. Figure Three presents these results, with DF* denoting the cases where the unit root was rejected by the $\tau_\tau$ test alone, and MTAR* denoting the equivalent case for the $\Phi_T^*$ test.

From inspection of Figure Three it can be seen that for breaks occurring early in the sample period, the $\tau_\tau$ test experiences more severe size distortion than the $\Phi_T^*$ test. Following the results of EG (1998), the DF test is known to have higher power than the MTAR test in the presence of symmetry. A practitioner applying these tests to an I(1) series which experiences an early break in sample would then erroneously conclude that the series was symmetric but stationary given the results of the test. False inferences would therefore be drawn concerning the integrated nature of the series.

However, a more interesting case emerges when considering the results for a break occurring late in the sample period. From the graph it can be seen that for late breaks the $\Phi_T^*$ test frequently rejects the unit root null when the $\tau_\tau$ test does not. In these circumstances, knowledge of the results of EG (1998), showing the MTAR test to be more powerful than the DF test in the presence of asymmetry of the MTAR form, would lead a practitioner to conclude that the series being examined is stationary with asymmetric adjustment. The inferences drawn would then be incorrect for two reasons, as the series would be deemed to be stationary and asymmetric when neither is true.

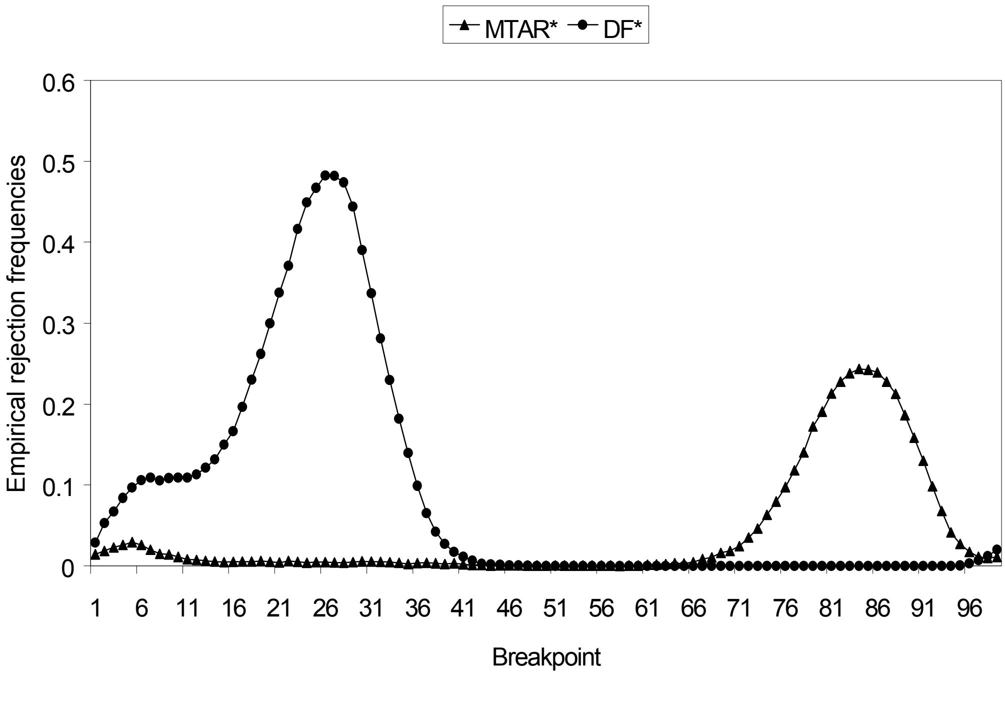


Figure Three: Conflicting testing outcomes (break in drift, $\alpha = 2$)

# 6 Conclusion

In this paper the size distortion of asymmetric unit root tests has been examined in the presence of breaks under the null hypothesis. Following the power analysis results of Enders (2001), attention focussed upon the behaviour of the original MTAR test rather than the consistent threshold equivalent. In the case of a levels break, the results obtained for the MTAR unit root test were broadly similar to those for the DF test presented by Leybourne *et al.* (1998), with distortion apparent only for breaks occurring early in the sample period. However when drift breaks were considered it was found that the MTAR test can experience severe size distortion when breaks occur both early and late in the sample period. A direct comparison of the DF and MTAR tests showed that for breaks in the drift parameter occurring late in the sample period, a practitioner would falsely conclude that an I(1) series was both stationary and symmetric. The results therefore suggest the MTAR test to be less robust to structural breaks than the familiar DF statistic, with the spurious rejection of both non-stationarity and symmetry possible.

# References

[1] Bai, J., Lumsdaine, R. and J. Stock (1998) 'Testing for and dating common breaks in multivariate time series', *Review of Economic Studies*, **65**, 395-432.

[2] Bai, J. and P. Perron (1998) 'Testing for and estimation of multiple structural changes', *Econometrica*, **66**, 47-79.

[3] Banerjee, A., Lumsdaine, R. and J. Stock (1992) 'Recursive and sequential tests of the unit root and trend break hypotheses: Theory and international evidence', *Journal of Business and Economic Statistics*, **10**, 271-287.

[4] Chan, K.S. (1993) 'Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model', *Annals of Statistics*, **21**, 520-533.

[5] Chan, K.S. and R.S. Tsay (1998) 'Limiting properties of the least squares estimator of a continuous threshold autoregressive model', *Biometrika*, **85**, 413–426.

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

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

[8] Dickey, D. and W. Fuller (1981) 'Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica*, **49**, 1057-1072.

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

[10] Enders, W. and C. W. J. Granger (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.

[11] Fuller, W. (1976) *Introduction to statistical time series*, New York: Wiley.

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

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

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

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

[16] Tong, H. (1990) *Non-linear time-series: a dynamical approach*, Oxford: Oxford University Press.

[17] Zivot and D. Andrews (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.