Finite-Sample Properties of the GLS-Based Dickey-Fuller Test in the Presence of Breaks in Innovation Variance

Steven Cook
University of Wales Swansea

Abstract: Using local-to-unity detrending, the GLS-based Dickey-Fuller test has been shown to possess higher power than other available unit root tests. As a result, application of this easily implemented test has increased in recent years. In the present study the finite-sample size and power of the GLS-based Dickey-Fuller test is examined in the presence of breaks in innovation variance under the null. In contrast to the original Dickey-Fuller test which has been shown to suffer severe distortion in such circumstances, the GLS-based test latter exhibits robustness to all but the most extreme breaks in variance. The results derived show the GLS-based test to be more robust to variance breaks than other modified Dickey-Fuller tests previously considered in the literature.

Zusammenfassung: Mittels Local-to-Unity Detrending wurde gezeigt, dass der GLS-basierte Dickey-Fuller Test eine größere Macht hat als andere verfügbare Unit-Root Tests. Deshalb kam es in den letzten Jahren zur vermehrten Anwendung dieses leicht zu implementierenden Tests. In der vorliegenden Studie wird unter der Nullhypothese für endliche Stichproben das Niveau und die Macht des GLS-basierten Dickey-Fuller Tests bei Einbrüchen der Innovationsvarianz untersucht. Im Gegensatz zum originalen Dickey-Fuller Test, welcher in derartigen Situationen eine starke Verzerrung erleidet, weist hierbei der GLS-basierte Test bis auf wenige Ausnahmen bei den stärksten Einbrüchen ein robustes Verhalten auf. Die erzielten Ergebnisse zeigen, dass der GLS-basierte Test auf Varianzbrüche robuster reagiert als andere modifizierte Dickey-Fuller Tests, welche bisher in der Literatur betrachtet wurden.

Keywords: Unit Roots, Variance Breaks, Local-to-Unity Detrending, Generalized Least Squares.

1 Introduction

It has long been recognised in the econometrics literature that the seminal Dickey-Fuller (DF) unit root test (Dickey and Fuller, 1979) possesses low power when applied to near-integrated, or near unit root, processes. Consequently, a number of modified DF tests with higher power have been proposed for practitioners to employ. Monte Carlo analysis of these tests has shown the modified DF test of Elliott et al. (1996) to possess the greatest power of all available unit root tests. The easily applied test of Elliott et al. (1996) employs local-to-unity detrending via generalised least squares estimation to produce a test with a local power function close to the local Gaussian power envelope. As a result of its high power and incorporation in software packages, application of this test, denoted as DF-GLS, has increased in recent years. Given the increasing popularity of the DF-GLS test amongst practitioners, investigation of its properties warrants close attention.

In this paper, the behaviour of the DF-GLS is examined in the presence of changes in innovation variance under the null. The recent research of Kim et al. (2002) has shown the DF test of Dickey and Fuller (1979) to suffer severe size distortion when applied to unit root processes experiencing a sudden decrease in innovation variance, with the unit root hypothesis spuriously rejected. However, while Cook (2002) and Cook (2003) have analysed the properties of a number of modified unit root tests in these circumstances, the properties of the DF-GLS test have not been considered in the literature to date. The objective of this paper is to therefore examine the finite-sample size and power of the DF-GLS test in the presence of decreases in innovation variance under the null.[1] It will therefore be examined how the DF-GLS performs and how this compares to previously derived results for the DF test and the variance break robust test of Kim et al. (2002).

2   Unit Root Testing in the Presence of Variance Breaks

2.1   Unit Root Tests

The behaviour of the DF τ_μ test when applied to unit root processes subject to a sudden decrease in innovation variance is examined in Kim et al. (2002).[2] The DF τ_μ test is given as the t-ratio for ϕ in the following regression:

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

To improve upon the known low power of the DF test, Elliott et al. (1996) proposes local-to-unity detrending using generalized least squares. This increased power is examined in Burridge and Taylor (2000), with results presented which show it results from the distribution of the DF-GLS test shifting further towards the origin under the null than under the alternative, in comparison to the DF test. Following the notation of Ng and Perron (2001), the DF-GLS equivalent of the τ_μ test is derived as follows. Given any series $\{x_t\}_{t=0}^T$, we define $(x_0^\alpha, x_t^\alpha) \equiv (x_0, (1 - \alpha \mathsf{L})x_t)$, $t = 1, \ldots, T$ where $\alpha = 1 + \bar{c}/T$. The revised, GLS-detrended version of y_t, denoted as $\tilde{y}_t$, is then

$$\tilde{y}_t \equiv y_t - \widehat{\psi}' z_t \,,$$

where z_t denotes the deterministic terms employed. In this case, with an intercept only considered, $z_t = 1$. The estimator $\widehat{\psi}$ is given as the value minimising $S(\alpha, \psi) = (y^\alpha - \widehat{\psi}' z^\alpha)'(y^\alpha - \widehat{\psi}' z^\alpha)$. For the intercept only model, the value $\bar{c} = -7$ is selected to ensure the limiting power function lies close to the local power envelope. With the locally detrended series $\tilde{y}_t$ calculated, the DF-GLS test is the t-ratio of β_0 in the regression

$$\Delta \tilde{y}_t = \beta_0 \tilde{y}_{t-1} + \varepsilon_t \tag{2}$$

---

[1]To illustrate the empirical relevance of such breaks in variance, Aggarwal et al. (1999) employ the Iterative Cumulative Sums of Squares (ICSS) algorithm of Inclan and Tiao (1994) to demonstrate that large changes in the innovation variance have recently occured in emerging stock markets.

[2]In Kim et al. (2002) it is shown that the size of the DF test is distorted by a *decrease* in innovation variance, but not an *increase*. As this property is found to hold for the DF-GLS test also, only decreases in innovation variance are considered.

with the resulting test statistic can be denoted as τ_μ^{GLS}. The appeal of this easily applied test to practitioners is obvious given its known increased power to reject the unit root hypothesis in the standard case of i.i.d. errors (see Elliott et al., 1996 for a power comparison of the τ_μ and τ_μ^{GLS} tests).

However, while the τ_μ^{GLS} test has not been examined in the presence of variance breaks, the properties of other (lower powered) modified tests have been considered in these circumstances. Although not considered explicitly in the present study, these modified tests are referred to later in terms of their properties in the presence variance breaks, and so are described here. The three tests considered are the weighted symmetric DF test of Park and Fuller (1995), the maximum DF test of Leybourne (1995) and the recursively mean adjusted test of Shin and So (2001). The weighted symmetric DF test of Park and Fuller (1995) results from the application of a double length regression, with the weighted symmetric estimator of the autoregressive parameter, denoted as $\widehat{\rho}_{ws}$, given as the value minimising

$$Q_{ws}(\rho) = \sum w_t(y_t - \rho y_{t-1})^2 + \sum (1 - w_t)(y_t - \rho y_{t+1})^2 \,, \tag{3}$$

where $w_t = (t-1)/T$. The unit root hypothesis is then tested using via the statistic τ_μ^{ws}

$$\tau_\mu^{ws} = \sigma_{ws}^{-1}(\widehat{\rho}_{ws} - 1)\left(\sum_{t=2}^{T-1} y_t^2 + T^{-1}\sum_{t=1}^{T} y_t^2\right)^{1/2}, \tag{4}$$

where $\sigma_{ws}^2 = (T-2)^{-1}Q_{ws}(\widehat{\rho}_{ws})$.

A further modified DF test proposed to increase power to reject the unit root hypothesis is the maximum DF test of Leybourne (1995). This test requires joint application of forward and reverse regressions. Given a series of interest $\{y_t\}_{t=0}^T$, the DF test of Leybourne (1995) is applied to both $\{y_t\}$ and $\{y_t^*\}$, where $z_t = y_{T-t}^*$ for $t = 0,\ldots,T$. The maximum DF test, denoted as τ_μ^{max}, is then simply the maximum (less negative) of the two test statistics obtained.

A final modified test previously considered is the recursively mean-adjusted DF test of Shin and So (2001). As Shin and So (2001) note, the use of mean-adjusted observations $(y_t - \overline{y})$ in the following regression results in correlation between the regressor $(y_{t-1} - \overline{y})$ and the error (ϵ_t)

$$y_t - \overline{y} = \gamma(y_{t-1} - \overline{y}) + \epsilon_t \,. \tag{5}$$

The resulting bias of the ordinary least squares estimator $\widehat{\gamma}$ has been calculated by Shaman and Stine (1988) and by Tanaka (1984) as

$$\mathrm{E}(\widehat{\gamma} - \gamma) = -T^{-1}(1 + 3\rho) + o(T^{-1}) \,. \tag{6}$$

To overcome the problems associated with this correlation, Shin and So (2001) propose the use of recursively mean-adjusted observations, with the recursive mean $(\overline{y}_t)$ calculated as

$$\overline{y}_t = t^{-1}\sum_{i=1}^{t} y_i \,. \tag{7}$$

The recursively mean-adjusted DF test, denoted as τ_μ^{rec}, is then given as the t-test of $\gamma = 1$ in the following regression

$$y_t - \overline{y}_{t-1} = \gamma(y_{t-1} - \overline{y}_{t-1}) + \epsilon_t \,. \tag{8}$$

2.2   Monte Carlo Design

To examine the properties of τ_μ and τ_μ^{GLS} in the presence of breaks in innovation variance, the following data generation process (DGP) is employed for the Monte Carlo simulation analysis

$$y_t = \rho y_{t-1} + \varepsilon_t \qquad t = 1, \ldots, T \tag{9}$$

$$\varepsilon_t = \sigma_t \eta_t \tag{10}$$

$$\eta_t \overset{iid}{\sim} \mathrm{N}(0,1) \tag{11}$$

$$\sigma_t = \begin{cases} \sigma_1 \text{ for } t \leq T_B \\ \sigma_2 \text{ for } t > T_B \end{cases} \tag{12}$$

where T_B represents the breakpoint after which an abrupt change in the innovation variance occurs. The innovation series $\{\eta_t\}$ is generated using pseudo iid $\mathrm{N}(0,1)$ random numbers from the RNDNS procedure in the GAUSS, with the initial value (y_0) set equal to zero. All experiments are performed over 10,000 replications using a sample size of $T = 100$, with a further initial 100 observations created and discarded.

Following Kim et al. (2002), the break ratio $\delta = \sigma_2/\sigma_1$ is assigned the values $\delta \in \{0.25, 0.4, 0.6, 0.8\}$. However, in contrast to Kim et al. (2002), results are presented for all possible breakpoints, $T_B \in \{1, 2, \ldots, 99\}$, rather than a limited number of points. To analyse the *size* properties of τ_μ and τ_μ^{GLS}, $\rho = 1$ is imposed in the DGP given by (3)-(6), while *power* is examined by imposing $\rho = 0.9$. Empirical rejection frequencies are calculated at the 5% level of significance using the critical values of Ng and Perron (2001) and Fuller (1996) for τ_μ and τ_μ^{GLS} respectively.

2.3   Monte Carlo Results

To ease interpretation, the results derived using the above experimental design are depicted graphically, with τ_μ and τ_μ^{GLS} denoted as DF and DF-GLS respectively, and the size of the break considered reported in parentheses. In Figure 1 the empirical sizes of the tests are reported for the larger variance breaks $(\delta = 0.25, 0.4)$. The known severe size distortion of τ_μ is apparent for both break sizes, particularly for breaks occurring in the early part of the sample. For the larger break $(\delta = 0.25)$ the empirical size of τ_μ reaches a maximum value of 41.35% when the break occurs at $T_B = 13$, while for $\delta = 0.4$ a maximum rejection frequency of 23.88% is recorded for $T_B = 14$. While distortion is maximized at these particular breakpoints, it is also apparent for breaks occurring at all other points, apart from the final few observations. In contrast, τ_μ^{GLS} exhibits far less size distortion, with the maximum empirical sizes of 10.32% and 7.46% obtained for $\delta = 0.25$ and 0.4. Note that this distortion for early breaks disappears when the break is imposed later in the sample. In Figure 2, the empirical sizes of the tests are presented for the two smaller breaks $(\delta = 0.6, 0.8)$. It can be seen that τ_μ still exhibits size distortion for early breaks, although this is now much reduced, the maximum rejection frequencies being 12.23% and 7.42% for $\delta = 0.6$ and 0.8 respectively. For these smaller breaks there is little evidence of size distortion for τ_μ^{GLS}, wherever the break is imposed, the maximum rejection frequencies being 5.76% and 5.18% for $\delta = 0.6$ and 0.8 respectively.
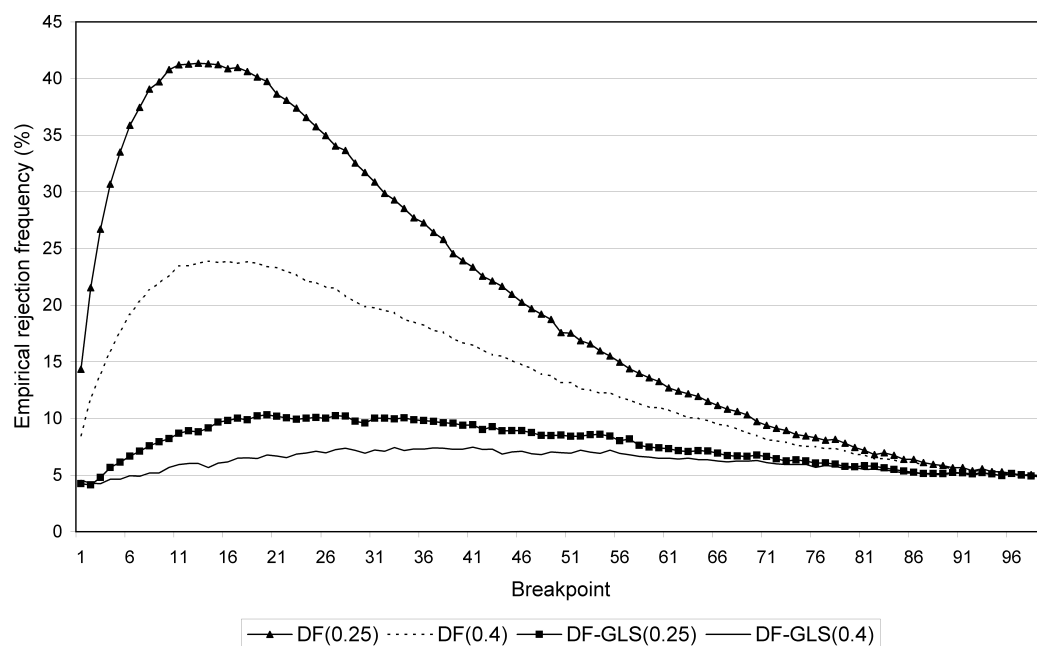
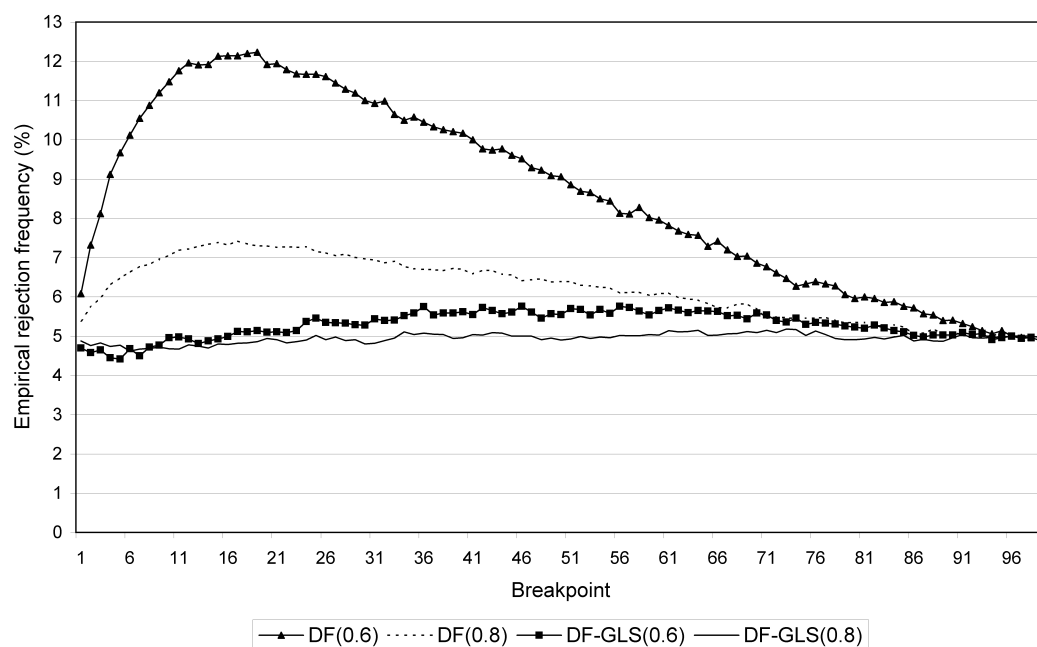Figure 1: Size of τ_μ and τ_μ^{GLS} tests ($\delta = 0.25, 0.4$).

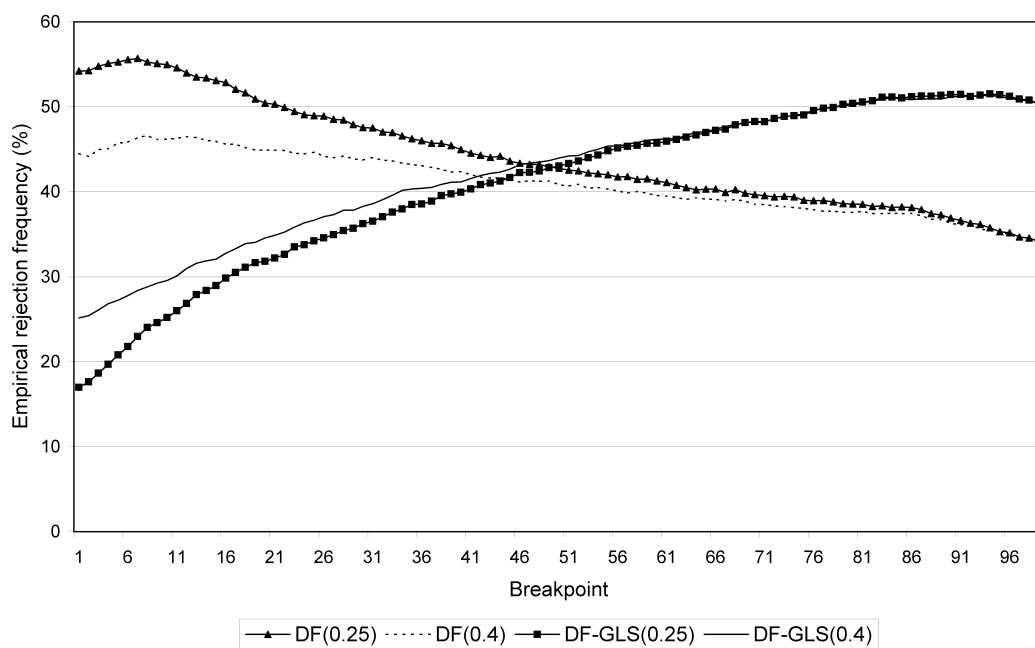Figure 2: Size of τ_μ and τ_μ^{GLS} tests ($\delta = 0.6, 0.8$).

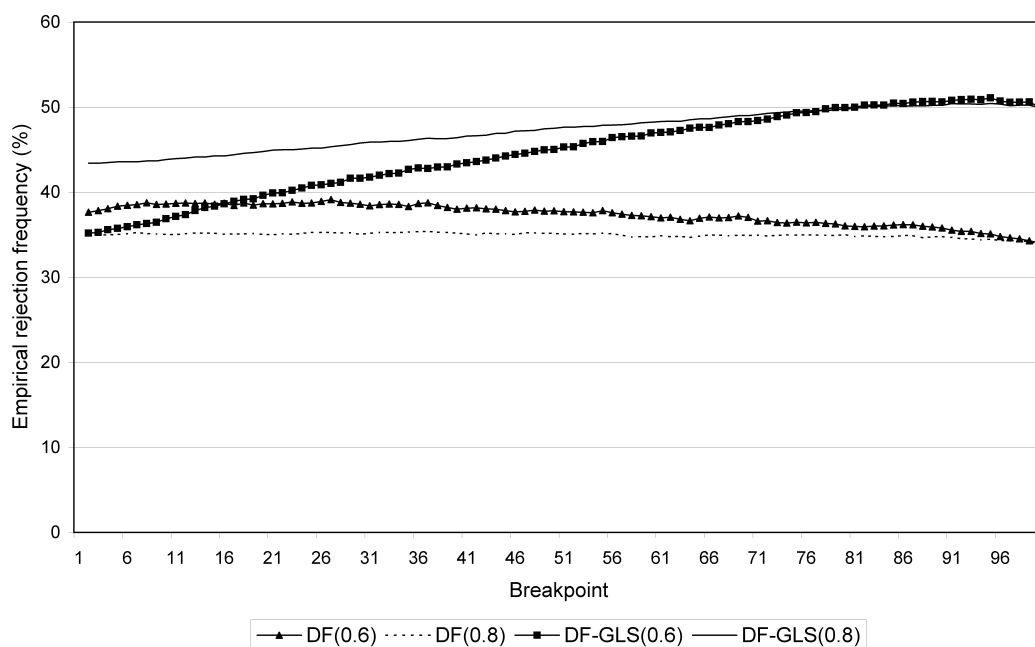Figure 3: Power of τ_μ and τ_μ^{GLS} tests ($\delta = 0.25, 0.4$).

Figure 4: Power of τ_μ and τ_μ^{GLS} tests ($\delta = 0.6, 0.8$).

Power calculations for the two tests are presented in Figures 3 and 4. The results for the larger breaks in Figure 3 show the power of τ_μ to be inflated by the previously noted size distortion for breaks early in the sample, while the known high power of τ_μ^{GLS} is reduced in the presence of early variance breaks. As expected, these characteristics are less noticeable for $\delta = 0.4$ than $\delta = 0.25$. However, as the timing of the break is delayed, the powers of the tests approach the levels noted previously in the literature for the no-break case, with τ_μ^{GLS} substantially more powerful than τ_μ. For the larger breaks of $\delta = 0.6$ and 0.8 depicted in Figure 4, it is clear that variance breaks have little impact upon the tests, apart perhaps for τ_μ^{GLS} when a break of $\delta = 0.6$ is imposed early in the sample.

Considering the results for size and power jointly, τ_μ^{GLS} is far more robust to breaks in innovation variance than τ_μ. In contrast to τ_μ, there is little evidence of τ_μ^{GLS} experiencing size distortion, except for very large breaks occurring early in the sample period. It should be noted that the value $\delta = 0.25$ which results in size distortion for τ_μ^{GLS} over a range of breakpoints is particularly extreme, corresponding to the innovation variance decreasing by a factor of 16. Similarly, the impact of innovation variance breaks upon the power of τ_μ^{GLS} is apparent for only a subset of the experiments considered, these again being where a very large break occurs early in the sample period.

2.4   A Comparison with the Variance Robust Test

The above results show the τ_μ^{GLS} test to be relatively robust in the presence of breaks in innovation variance. To consider how the test compares to the variance robust test (t_F) of Kim et al. (2002), Tables 1 and 2 present the size and power calculations for the τ_μ^{GLS}, τ_μ and t_F tests across a range of breakpoints and breaks sizes. Following the tabulated results of Kim et al. (2002), break sizes of $\delta = \{0.25, 0.4, 0.6, 0.8\}$ are considered at the breakpoints $T_B = \{20, 40, 60, 80\}$. As might be expected, the results in Table 1 show that the t_F test has the best size properties in the presence of the largest break in variance considered. However, for more moderate breaks, the τ_μ^{GLS} test actually outperforms the asymptotically justified t_F test. Considering the power results presented in Table 2, it can be seen that the power of the τ_μ^{GLS} test exceeds that of the t_F test in virtually all cases considered. It can therefore be argued that in many circumstances the τ_μ^{GLS} may be preferred to the t_F test in the presence of breaks in variance, unless the breaks are extremely large $(\delta = 0.25)$.

3   Conclusion

In this letter the properties of the GLS-based Dickey-Fuller test of have been examined in the presence of breaks in innovation variance. In contrast to the Dickey-Fuller test which can exhibit severe size distortion when applied to a unit root process experiencing a sudden decrease in innovation variance, the size of the DF-GLS test is found to be substantially more robust. Additionally, the results of power calculations show the DF-GLS test to only experience reduced power when very large breaks in variance occur early in the sample period. For more moderate and realistic break sizes, the properties of the

Table 1: Empirical sizes of the τ_μ^{GLS}, τ_μ and t_F tests ($T = 100$).

T_B		δ			
		0.80	0.60	0.40	0.25
20	τ_μ	7.3	11.9	23.4	39.7
	τ_μ^{GLS}	**4.9**	**5.1**	6.7	10.3
	τ_F	6.1	6.8	**6.3**	**6.3**
40	τ_μ	6.7	10.2	16.7	23.9
	τ_μ^{GLS}	**5.0**	5.6	7.3	9.4
	τ_F	5.4	**5.2**	**5.5**	**5.4**
60	τ_μ	6.1	8.0	11.0	13.3
	τ_μ^{GLS}	**5.0**	5.7	**6.5**	7.4
	τ_F	4.7	**4.7**	**5.5**	**5.6**
80	τ_μ	5.4	6.0	6.8	7.4
	τ_μ^{GLS}	**4.9**	**5.2**	5.7	5.7
	τ_F	4.5	4.3	**5.0**	**5.3**

Entries in bold indicate the test statistic with an empirical size closest to the nominal level of 5% for the combination of break size and breakpoint considered.

Table 2: Empirical power of the τ_μ^{GLS}, τ_μ and t_F tests ($T = 100$).

T_B		δ			
		0.80	0.60	0.40	0.25
20	τ_μ	35.0	38.6	44.9	50.4
	τ_μ^{GLS}	45.0	39.9	34.6	31.8
	τ_F	26.7	27.9	34.0	46.5
40	τ_μ	35.1	38.1	42.4	45.0
	τ_μ^{GLS}	46.6	43.5	41.2	40.0
	τ_F	24.3	24.9	28.0	38.7
60	τ_μ	34.9	37.0	39.6	41.3
	τ_μ^{GLS}	48.3	47.0	46.2	45.7
	τ_F	24.4	22.4	25.8	35.0
80	τ_μ	34.8	36.0	37.6	38.6
	τ_μ^{GLS}	49.9	50.0	50.2	50.4
	τ_F	25.3	24.2	26.4	30.8

DF-GLS test are even superior to those of the variance robust test of Kim et al. (2002). Given the extreme nature of the larger breaks considered here and by Kim et al. (2002), perhaps more importance should be attached to these results for the more moderate breaks

where the DF-GLS performs well. The observed robustness of the GLS test in these circumstances, allied to its known high power in the absence of breaks, further emphasizes its value to practitioners. A comparison with the findings of Cook (2003) shows the easily applied GLS test to have similar properties to the recursively mean-adjusted test of Shin and So (2001), but outperform the weighted symmetric test of Park and Fuller (1995). A further comparison with the results of Cook (2002), shows the DF-GLS test to outperform the maximum DF test of Leybourne (1995) in the presence of various breaks. The appeal of the GLS test is further emphasised as unlike the recursively mean-adjusted test, its specification can be extended to include a deterministic trend.

Acknowledgement

References

Aggarwal, R., Inclan, C., and Leal, R. (1999). Volatility in emerging stock markets. *Journal of Financial and Quantitative Analysis*, *34*, 33-55.

Burridge, P., and Taylor, A. (2000). On the power of GLS-type unit root tests. *Oxford Bulletin of Economics and Statistics*, *62*, 633-645.

Cook, S. (2002). Unit root testing in the presence of innovation variance breaks: A simple solution with increased power. *Journal of Applied Mathematics*, *5*, 233-240.

Cook, S. (2003). Size and power properties of powerful unit root tests in the presence of variance breaks. *Physica A*, *317*, 432-448.

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

Elliott, G., Rothenberg, T., and Stock, J. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, *64*, 813-836.

Fuller, W. (1996). *Introduction to statistical time series* (second ed.). New York: Wiley.

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. (2002). Unit root tests with a break in variance. *Journal of Econometrics*, *109*, 365-387.

Leybourne, S. (1995). Testing for unit roots using forward and reverse Dickey-Fuller regressions. *Oxford Bulletin of Economics and Statistics*, *57*, 559-571.

Ng, S., and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, *69*, 1519-1554.

Park, H., and Fuller, W. (1995). Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis*, *16*, 415-429.

Shaman, P., and Stine, R. (1988). The bias of autoregressive coefficient estimators. *Journal of the American Statistical Association*, *83*, 842-848.

Shin, D., and So, B. (2001). Recursive mean adjustment for unit root tests. *Journal of Time Series Analysis*, *22*, 595-612.

Tanaka, K. (1984). An asymptotic expansion associated with maximum likelihood estimators in ARMA models. *Journal of the Royal Statistical Society, B*, *46*, 58-67.

Author's address:

Dr. Steven Cook
Department of Economics
University of Wales Swansea
Singleton Park
Swansea, SA2 8PP

Tel. (01792) 602106
E-mail: s.cook@swan.ac.at