Finite sample effects of pure seasonal mean shifts on Dickey-Fuller tests: a simulation study *

Artur C. B. da Silva Lopes †
CEMAPRE and Instituto Superior de Economia e Gestão (ISEG–TULisbon)

This version: November 10, 2007

Abstract

In this paper, it is demonstrated by simulation that, contrary to a widely held belief, pure seasonal mean shifts – i.e., seasonal structural breaks which affect only the seasonal cycle – really do matter for Dickey-Fuller long-run unit root tests. Both size and power properties are affected by such breaks but using the t-sig method for lag selection induces a stabilizing effect. Although most results are reassuring when the t-sig method is used, some concern with this type of breaks cannot be disregarded.

Further evidence on the poor performance of the t-sig method for quarterly time series in standard (no break) cases is also presented.

Keywords: unit roots; seasonality; Dickey-Fuller tests; structural breaks.

JEL Classification: C22, C52

*I am grateful to Ann Henshall, to the participants of the Growth and Business Cycles in Theory and Practice Conference (Manchester) and to two anonymous referees for helpful comments and suggestions. Financial support from Fundação para a Ciência e Tecnologia (FCT) through Programa POCTI, partially funded by FEDER, is also gratefully acknowledged.

†Address: ISEG-UTL, Rua do Quelhas 6, Gab. 315, 1200 Lisboa, Portugal. E-mail: asl@iseg.utl.pt. Fax: +351-21-3922781. Phone: +351-21-3922796.

1 Introduction

In an attempt to improve the quality of inferences produced by Dickey-Fuller (1979) (DF) tests, these are very often applied to infra-annual, seasonally unadjusted or raw data. Power properties are particularly envisaged with such larger sample sizes. On the other hand, a recent strand of research has been investigating the effects of breaks in the seasonal pattern of economic time series on tests for seasonal unit roots; see, e.g., Smith and Otero (1997), Franses and Vogelsang (1998), Balcombe (1999), Lopes (2001), Montañés and Sansó (2001), Harvey et al. (2002), Hassler and Rodrigues (2004), and Lopes and Montañés (2005).

However, it has been assumed more or less explicitly that *pure seasonal mean shifts* (PSMSs) — i.e., structural breaks in the deterministic seasonal cycle that leave the level and the slope of the trend function unchanged — should not affect the properties of tests designed for analysing the long-run properties of the series. This assumption is rooted in the traditional trend-cycle-seasonal decomposition approach, which has evolved towards "conventional wisdom", common belief or intuition. According to this intuition, such breaks change the behaviour of the series at seasonal frequencies only, and hence only seasonal unit root tests are expected to have their performance disturbed.

In this paper, it is shown that this intuition is incorrect. The key for this argument builds on Lopes (2006) and it is based on the perspective of viewing seasonal mean shifts as neglected systematic additive outliers (AOs). This sheds a different light over the subject, making clear that the performance of long-run (A)DF unit root tests is also disturbed by such mere seasonal breaks.

For the simple AR(1) model, the shift of the DF distribution to the left induced by PSMSs imparts straightforward effects: a spurious rejection problem under the null hypothesis, and a corresponding improvement of the power properties under the alternative. The intuition is simple: in both cases PSMSs may be seen as highly transitory shocks and hence weaken (strengthen) the evidence for (against) the null hypothesis. However, the analysis is further complicated when more realistic processes and testing strategies are entertained, as is usually the case in practice. In particular, data dependent methods for lag selection are known to induce a counter-effect to that of AOs. Hence, evaluating the final effects in small samples requires a somewhat extensive Monte Carlo study. Although most results are reassuring, some concern with PSMSs cannot be disregarded.

The remainder of this paper is organized as follows. The next section introduces the model of pure seasonal mean shifts and develops the main ideas. Section 3 contains the results of the Monte Carlo study, both for the $I(1)$ and $I(0)$ cases. Section 4 concludes.

2 Dickey-Fuller tests and pure seasonal mean shifts

Consider the case of quarterly, seasonally unadjusted and trending data, and the standard (non-augmented) DF test regression for a unit root:

$$\Delta y_t = \sum_{j=1}^{4} \hat{\theta}_j \, D_{tj} + \hat{\beta}\, t + \hat{\phi}\, y_{t-1} + \hat{u}_t, \quad t = 2, 3, ..., T, \tag{1}$$

where D_{tj} $(j = 1, 2, 3, 4)$ denote the usual seasonal dummy variables, estimated by OLS to produce the usual t-ratio, $\hat{\phi}/\hat{\sigma}_{\hat{\phi}}$. Suppose that the data generating process (DGP) is

$$y_t = \alpha + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] \, D_{tj} + \beta\, t + \epsilon_t, \tag{2}$$

where $\sum_{j=1}^{4} \gamma_j = 0$, $I(t > \tau)$ is an indicator function equal to 1 when $t > \tau = \lambda T$ and 0 otherwise, λ denoting the fraction break parameter, δ_j $(j = 1, ..., 4)$ represent the parameters of the seasonal mean shifts, and ϵ_t is iid $N(0, \sigma^2)$. Hence, the seasonal cycle is represented by the γ_j parameters until the time of the break and by $\gamma_j + \delta_j$ after that time.

When $\sum_{j=1}^{4} \delta_j \neq 0$, the neglected seasonal break changes the level of the series too. From the work of Perron (1989) and the research that followed, it is known that, despite remaining consistent, standard DF tests may become rather powerless in small samples.

On the other hand, in the case where the DGP is the unit root process,

$$\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] \, D_{tj} + \epsilon_t, \tag{3}$$

with y_0 fixed, when $\sum_{j=1}^{4}\delta_j \neq 0$, the "converse Perron phenomenon" emerges: Leybourne *et al.* (1998) and Leybourne and Newbold (2000) have shown that DF tests that disregard the break may produce spurious evidence against the (true) null hypothesis.

Consider now that $\sum_{j=1}^{4}\delta_j = 0$, i.e., the structural break is constrained to affect the deterministic seasonal cycle only, the level and the slope of the trend remaining unchanged. In this paper, this case is termed as a *pure seasonal mean shifts* (PSMSs) case. It is commonly believed that neglecting such a break produces no effect on the asymptotic and finite sample distribution of DF statistics. The intuition is that since only the seasonal cycle changes, contamination remains confined to the seasonal frequencies, and hence, as it does not go through the long run or zero frequency properties of the data, the behaviour of DF statistics should remain unchanged. This reasoning, however, neglects the crucial importance of correctly accounting for deterministic terms in DF regressions.

A similar problem occurs when deterministic seasonality is present in the DGP but it is neglected in the testing strategy. Based in the work by Franses and Haldrup (1994), Lopes (2006) [1] addresses the consequences of neglecting deterministic seasonality in DF regressions considering the effect of the omitted usual set of seasonal dummy regressors as systematic additive outliers (AOs). The same approach is used here, extending it to the case of PSMSs [2]. Considering these as systematic AOs opens a very different and insightful perspective: neglected PSMSs induce a negative "MA-like" component in the errors of the test regression whose effects on DF tests are well known since the work of Schwert (1989). Intuitively, the mechanism is also simple: neglected PSMSs spuriously inflate the transitory dynamics of the series, shifting the DF distribution

[1] Demetrescu and Hassler (2007) address the same problem following a different route.

[2] The connection between omitted seasonal dummies and systematic AOs in DF regressions was established in Lopes (2006) for a particular case of the seasonal pattern to facilitate deriving the asymptotic distribution of DF statistics. The same connection can be easily established here for the case where the (pure) seasonal break is of the type $\sum_{j=1}^{4}\delta_j = \sum_{j=1}^{4}(-1)^j\delta = 0$ (see Lopes, 2006, pp. 168-9); i.e., neglecting deterministic terms unrelated with the break, equation (3) can be written using the framework of Franses and Haldrup (1994) for the post-break period. Hence, now the problem is further complicated due to the presence of a parameter representing the fraction of the sample where the break occurs (τ/T).

to the left, both under the null and the alternative hypothesis. In particular, under the null hypothesis, unmodeled PSMSs induce a spurious mean reverting behaviour into the deterministic component of the series. Therefore, spurious unit root rejections are predicted when the DGP is given by equation (3), and boosted power properties are expected when the DGP is a trend stationary process [3].

In practical terms, analysing the failure to account for PSMSs presents further difficulties. First, DGPs as simple as the ones of equations (2) and (3) are overly simplistic and are rarely (if ever) seriously considered in empirical work. Second and most important, even when the DGP coincides with (2) or (3), empirical work usually begins with an autoregressive model whose order is a function of T, and then proceeds using some rule for lag selection. That is, instead of (1) above, (A)DF statistics usually result from equations such as

$$\Delta y_t = \sum_{j=1}^{4} \hat{\theta}_j \, D_{tj} + \hat{\beta}\, t + \hat{\phi}\, y_{t-1} + \sum_{i=1}^{\hat{k}} \hat{\psi}_i \, \Delta y_{t-i} + \hat{e}_t, \tag{4}$$

where $\hat{k}$ is selected using some data dependent method.

Previous research has found that, at least in small samples, this procedure is liable to produce further distortions, albeit in the opposite direction of the one resulting from PSMSs. In particular, the popular general-to-specific (GS) t-sig method tries to capture the "MA-like" effect of AOs through inflated lag augmentation; see, e.g., Perron and Rodríguez (2003) and Lopes (2006). As a result, since the distribution tends to shift to the right, a kind of counter-effect is observed, both in terms of size and power. Determining the final, combined outcome in empirically relevant cases and sample sizes demands an extensive simulation study.

[3] See Franses and Haldrup (1994) for the asymptotic and numerical analysis of the effect of AOs on DF tests under the null hypothesis. For the alternative, besides Lopes (2006), the closest study addressing this issue appears to be Nelson *et al.* (2001), containing numerical evidence for a case where a Markov regime switching process is assumed for the transitory component of the series (see their subsection 1.2).

3 Simulation results

To simplify the analysis, a single break parameter, δ, is considered, and the results for the following cases are presented:

$$\begin{array}{ll} \text{case A} & \delta = \delta_1 = -\delta_2,\ \delta_3 = \delta_4 = 0; \\ \text{case B} & \delta = \delta_1 = -\delta_3,\ \delta_2 = \delta_4 = 0; \\ \text{case C} & \delta = \delta_1 = \delta_3 = -\delta_2 = -\delta_4, \text{ and} \\ \text{case D} & \delta = \delta_1 = \delta_2 = -\delta_3 = -\delta_4. \end{array}$$

For instance, for case A and when $\delta > 0$, a *boom* in the first quarter is balanced by a *crash* in the second quarter. Two other cases were considered ($\delta = \delta_1$, $-\delta/3 = \delta_2 = \delta_3 = \delta_4$ and $\delta = \delta_1$, $-\delta/2 = \delta_2 = \delta_3$, $\delta_4 = 0$) but their evidence adds little to the analysis. For the cases reported the relevant magnitude is the "standardized break", δ/σ_ϵ, i. e., the relative size of the break. Hence, with no loss of generality, $\sigma_\epsilon = 1$ and $\delta = 0$ (no break), 1, 3, and 5. As usual, the study concentrates on frequency rejections of nominal 5% level tests, the small samples critical values taken from MacKinnon (1991).

Unless stated otherwise, the reported results concern a break in the middle of the sample, i.e., $\lambda = 0.5$. For each experiment $10,000$ replications were generated using TSP 4.5. The sample sizes considered are for $T = 48$, 96, and 160. For the ADF tests, the GS t-sig procedure is initiated with $k_{max} = 4$, 8, and 12, respectively, and the reduction tests are based on a 5% asymptotic level critical region.

3.1 The $I(1)$ case: size

Table 1 contains the results for the case when the DGP is given by equation (3). While the left panel addresses the case for DF (non-augmented) tests [equation (1)], the right panel contains the results for ADF tests [equation (4)], $\bar{\hat{k}}$ denoting the average fitted lag length.

Table 1 around here

As expected, for DF tests a clear picture of size distortions emerges; the over-rejections grow with δ (and T) and become rather dramatic when $\delta = 5$.

Additional (unreported) simulations for all types of breaks and for the cases when $T = 96$, 160 and $\delta = 3$ reveal that spurious rejections are more severe for breaks located around the middle of the sample (i.e., for $\lambda \in [0.45; 0.55]$).

A quite different picture is observed for ADF tests. The GS t-sig method clearly helps in restoring size close to its nominal level, particularly for large δ (and T). This occurs because, as expected, the increase in δ is generally associated with larger lag lengths. Hence, although size distortions still subsist, with the exception of a few cases when $T = 48$ only, in general they seem tolerable for the sample sizes analysed.

Table 2 contains the size estimates for three additional, more realistic and demanding DGPs:

$$\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j]\, D_{tj} + \epsilon_t - 0.8\epsilon_{t-1}, \tag{5}$$

$$\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j]\, D_{tj} + \epsilon_t - 0.8\epsilon_{t-4}, \tag{6}$$

and

$$(1 + 0.9L)(1 + 0.4L^2)\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j]\, D_{tj} + \epsilon_t. \tag{7}$$

While the DGP from equation (5) is well known since the work by Schwert (1989), the one of equation (6) represents its traditional "seasonal twin". The DGP of equation (7) is purely autoregressive, i.e., it is less demanding than the previous two. Its interest derives from the fact that it is a near-semiannual unit root case, i.e., it allows some instability in the seasonal pattern involving adjacent quarters.

Table 2 around here

In order to facilitate the comparison with the frequency rejections for the no break case ($\delta = 0$), these are presented in bold. The following main conclusions may be drawn:

i) a generalized picture of size distortions emerges, even for the no-break case and particularly for the DGPs containing negative MA terms. This results from the poor performance of the t-sig method in small samples when deterministic regressors are needed to achieve similarity of DF tests; see Taylor (2000) and Lopes (2006).

ii) Somewhat unexpectedly, the presence of the break increases the frequency of spurious rejections only in about 1/3 of the cases. That is, the most frequent outcome of PSMSs is to improve, not to worsen, the size performance of ADF tests. This results from the lag augmentation inflation effect, which is clearly visible. A quite interesting, although extreme, example occurs for the DGP of equation (5) when $T = 48$: when there is no break the estimated size is 87.9% and $\bar{\hat{k}} = 0.53$; when $\delta = 5$ the estimated size is reduced to only 0.6% and $\bar{\hat{k}} = 2.49$.

iii) Most of the cases where size behaviour deteriorates occur with the DGP of equation (6). This is worrying because the presence of negative seasonal MA error terms is frequently reported in empirical studies; see, e.g., Ghysels *et al.* (1994).

3.2 The $I(0)$ case: power

The power analysis begins with the simple DGP of equation (1), with $\beta = 0.05$, $\rho \equiv \phi = 0.90$ and $u_t \equiv \epsilon_t \sim$ iid $N(0,\, 1)$. The results are presented in table 3.

Table 3 around here

As predicted, a boosted power performance is now observed for the DF test. For the more realistic setting of ADF tests, however, power reductions are observed in 50% of the cases, and these are more frequent and significant when $T = 160$, as k_{max} is permitted to attain 12 and a very large lag truncation parameter is usually selected. Gains in power tend to be relatively large only when $T = 48$.

A similar unclear picture is observed when the following DGPs are considered:

$$y_t = \alpha + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j]\, D_{tj} + 0.05\, t + 0.90\, y_{t-4} + \epsilon_t, \qquad (8)$$

$$(1-0.9L)(1+0.9L)(1+0.4L^2)y_t = \alpha + [\sum_{j=1}^{4}\gamma_j + I(t>\tau)\sum_{j=1}^{4}\delta_j]\, D_{tj} + 0.05\,t + \epsilon_t, \text{ i.e.,}$$

$$y_t = \alpha + [\sum_{j=1}^{4}\gamma_j + I(t>\tau)\sum_{j=1}^{4}\delta_j]\, D_{tj} + 0.05\,t + 0.41\,y_{t-2} + 0.324\,y_{t-4} + \epsilon_t, \quad (9)$$

and

$$y_t = \alpha + [\sum_{j=1}^{4}\gamma_j + I(t>\tau)\sum_{j=1}^{4}\delta_j]\, D_{tj} + 0.05\,t + 0.90\,y_{t-4} + \epsilon_t - 0.4\epsilon_{t-4}. \quad (10)$$

The costs of the generally satisfactory size performance, achieved through liberal lag lengths, now emerge from table 4:

i) power gains resulting from the additional "transitory dynamics" tend to evaporate, and in about 50% of the cases, power is actually reduced;

ii) in particular, large power reductions are observed for the DGP of equation (9), which acts as the stationary version of the one from equation (7);

iii) significant power gains tend to occur only when $T = 48$, i.e., when $\hat{k}$ is constrained to be low.

Table 4 around here

4 Concluding remarks

The widely held belief that pure seasonal mean shifts do not matter for Dickey-Fuller tests is incorrect. When they are not accommodated, transitory fluctuations in the errors of the test regressions are spuriously inflated, distorting inference not only about seasonal cycles but also about the long-run properties of economic time series.

In the AR(1) case and for the non-augmented version of the test, both the deterioration in size behaviour and the jump in the power function can be quite dramatic. In more realistic settings the consequences are usually milder and uncertain due to the stabilizing, inflationary lag augmentation, effect produced by the general-to-specific t-sig procedure. However, relying on this to cope with the problem can also be hazardous. In some cases where a unit root is present the frequency of spurious rejections may increase significantly. Under the alternative, for some cases power may be reduced. A method designed to detect such seasonal breaks and to account for them in the testing strategy is called for.

On the other hand, although the t-sig method can be very helpful in redressing size behaviour in the presence of breaks, additional evidence on its poor performance in small samples for standard (no break) cases was also obtained. A better method for lag selection is needed to complement Dickey-Fuller tests when they are used with quarterly data. Both issues are left for future research.

References

[1] Balcombe, K., 1999. Seasonal unit root tests with structural breaks in deterministic seasonality, *Oxford Bulletin of Economics and Statistics*, 61, 569-82.

[2] Demetrescu, M. and Hassler, U., 2007. Effects of neglected deterministic seasonality on unit root tests, *Statistical Papers*, 48, 385-402.

[3] Dickey, D. A. and Fuller, W. A., 1979. Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-31.

[4] Franses, P. H. and Haldrup, N., 1994. The effects of additive outliers on tests for unit roots and cointegration, *Journal of Business and Economic Statistics*, 12, 471-8.

[5] Franses, P. H. and Vogelsang, T., 1998. On seasonal cycles, unit roots, and mean shifts, *The Review of Economics and Statistics*, 80, 231-40.

[6] Ghysels, E., Lee, H. S. and Noh, J., 1994. Testing for unit roots in seasonal time series, *Journal of Econometrics*, 62, 415-442.

[7] Harvey, D. I., Leybourne, S. J. and Newbold, P., 2002. Seasonal unit root tests with seasonal mean shifts, *Economics Letters*, 76, 295-302.

[8] Hassler, U. and Rodrigues, P. M. M., 2004. Seasonal unit root tests under structural breaks, *Journal of Time Series Analysis*, 25, 33-53.

[9] Leybourne, S. J., Mills, T. C. 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.

[10] Leybourne, S. J. and Newbold, P., 2000. Behaviour of the standard and symmetric Dickey-Fuller-type tests when there is a break under the null hypothesis, *Econometrics Journal*, 3, 1-15.

[11] Lopes, A. C. B. da S., 2001. The robustness of tests for seasonal differencing to structural breaks, *Economics Letters*, 71, 173-9.

[12] Lopes, A. C. B. da S., 2006. Deterministic seasonality in Dickey-Fuller tests: should we care?, *Empirical Economics*, 31, 165-182.

[13] Lopes, A. C. B. da S. and Montañés, A., 2005. The behavior of HEGY tests for quarterly time series with seasonal mean shifts, *Econometric Reviews*, 24, 83-108.

[14] MacKinnon, J., 1991. Critical values for cointegration tests, *in* Engle, R. F. and Granger, C. W. J., eds., *Long-Run Economics Relationships*, Oxford University Press, 267-76.

[15] Montañés, A. and Sansó, A., 2001. The Dickey-Fuller test family and changes in the seasonal pattern, *Annales d'Économie et de Statistique*, 61, 73-90.

[16] Nelson, C. R., Piger, J. and Zivot, E., 2001. Markov regime switching and unit-root tests, *Journal of Business and Economic Statistics*, 19, 404-15.

[17] Perron, P., 1989. The great crash, the oil shock and the unit root hypothesis, *Econometrica*, 57, 1361-402.

[18] Perron, P. and Rodríguez, G., 2003. Searching for additive outliers in nonstationary time series, *Journal of Time Series Analysis*, 24, 193-220.

[19] Schwert, G. W. (1989). Tests for unit roots: a Monte Carlo investigation, *Journal of Business and Economic Statistics*, 7, 147-59.

[20] Smith, J. and Otero, J., 1997. Structural breaks and seasonal integration, *Economics Letters*, 56, 13-19.

[21] Taylor, A. M. R., 2000. The finite sample effects of deterministic variables on conventional methods of lag-selection in unit root tests, *Oxford Bulletin of Economics and Statistics*, 62, 293-304.

Table 1. Size estimates of nominal 5% level (A)DF tests for the random walk case

	DF tests			ADF tests		
		T			T	
δ	48	96	160	48 ($\bar{\hat{k}}$)	96 ($\bar{\hat{k}}$)	160 ($\bar{\hat{k}}$)
			case A: $\delta = \delta_1 = -\delta_2,\ \delta_3 = \delta_4 = 0$			
1	5.7	6.2	7.1	8.0 (0.58)	8.1 (2.01)	7.6 (4.72)
3	28.0	32.4	34.6	13.6 (2.92)	9.3 (1.89)	8.1 (9.83)
5	65.4	72.3	74.5	7.1 (3.83)	9.0 (1.85)	7.3 (9.41)
			case B: $\delta = \delta_1 = -\delta_3,\ \delta_2 = \delta_4 = 0$			
1	5.4	6.0	6.8	7.8 (0.57)	7.6 (2.15)	7.3 (4.91)
3	21.7	29.6	32.6	8.2 (2.52)	7.6 (5.52)	6.7 (8.69)
5	51.9	67.2	71.5	4.2 (2.82)	7.8 (5.26)	6.0 (7.94)
			case C: $\delta = \delta_1 = \delta_3 = -\delta_2 = -\delta_4$			
1	8.3	9.1	10.0	10.3 (1.04)	9.4 (3.40)	7.2 (6.60)
3	57.1	61.6	62.3	10.6 (2.14)	8.9 (4.18)	7.3 (6.52)
5	91.2	93.6	94.3	8.3 (1.95)	9.2 (3.46)	7.2 (5.42)
			case D: $\delta = \delta_1 = \delta_2 = -\delta_3 = -\delta_4$			
1	6.9	8.3	9.4	8.8 (0.82)	8.0 (3.29)	7.0 (7.00)
3	35.0	51.4	56.1	5.6 (2.89)	7.6 (5.47)	6.6 (8.29)
5	65.9	85.5	90.1	3.0 (2.74)	8.3 (4.78)	6.2 (6.97)

Table 2. Size estimates of nominal 5% level ADF tests for more realistic $I(1)$ processes

		T			T	
δ	48 ($\bar{\hat{k}}$)	96 ($\bar{\hat{k}}$)	160 ($\bar{\hat{k}}$)	48 ($\bar{\hat{k}}$)	96 ($\bar{\hat{k}}$)	160 ($\bar{\hat{k}}$)
DGP: $\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] D_{tj} + \epsilon_t - 0.8\epsilon_{t-1}$						
0	**87.9** (0.53)	**65.1** (2.07)	**46.1** (4.52)	**87.9** (0.53)	**65.1** (2.07)	**46.1** (4.52)
		Case A			Case B	
1	88.1 (0.47)	64.8 (2.02)	44.0 (4.63)	88.3 (0.45)	61.8 (2.11)	40.7 (4.88)
3	69.9 (1.12)	24.3 (4.54)	11.8 (8.45)	63.7 (1.31)	21.5 (4.40)	14.5 (7.67)
5	20.7 (2.65)	12.7 (5.54)	6.8 (9.07)	24.5 (2.26)	10.4 (4.60)	8.4 (7.67)
		Case C			Case D	
1	83.2 (0.56)	62.9 (2.34)	37.1 (5.23)	86.5 (0.47)	51.1 (2.59)	31.6 (5.74)
3	41.0 (1.36)	29.6 (3.89)	16.4 (6.81)	22.6 (2.16)	6.7 (4.67)	6.7 (7.80)
5	18.0 (1.42)	17.3 (3.70)	10.7 (6.08)	0.6 (2.49)	1.0 (4.54)	1.5 (7.03)
DGP: $\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] D_{tj} + \epsilon_t - 0.8\epsilon_{t-4}$						
0	**31.6** (2.80)	**25.6** (6.41)	**27.1** (10.12)	**31.6** (2.80)	**25.6** (6.41)	**27.1** (10.12)
		Case A			Case B	
1	39.4 (1.80)	51.7 (3.79)	68.0 (6.11)	38.6 (1.84)	49.4 (4.22)	51.9 (7.57)
3	56.5 (0.85)	50.0 (6.99)	41.6 (11.66)	51.1 (1.04)	23.8 (6.19)	23.3 (10.32)
5	55.6 (2.91)	37.8 (7.72)	37.1 (11.33)	21.7 (2.21)	16.1 (6.33)	18.6 (9.90)
		Case C			Case D	
1	49.2 (1.79)	54.7 (5.46)	48.8 (9.59)	39.9 (1.27)	57.3 (3.95)	32.1 (9.23)
3	56.3 (2.03)	46.8 (5.89)	43.9 (9.39)	29.5 (1.95)	13.8 (6.37)	14.2 (10.12)
5	49.6 (2.03)	36.1 (5.62)	41.6 (8.52)	7.4 (2.40)	5.5 (6.22)	7.0 (9.39)
DGP: $(1 + 0.9L)(1 + 0.4L^2)\Delta y_t = \beta + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] D_{tj} + \epsilon_t$						
0	**29.4** (1.25)	**13.9** (3.27)	**7.1** (4.89)	**29.4** (1.25)	**13.9** (3.27)	**7.1** (4.89)
		Case A			Case B	
1	27.1 (1.58)	7.7 (3.50)	5.9 (5.06)	32.7 (1.62)	7.6 (3.66)	5.9 (5.50)
3	8.8 (2.91)	6.6 (4.41)	7.1 (7.23)	6.6 (2.93)	6.7 (4.44)	7.1 (7.38)
5	5.9 (3.16)	7.0 (4.93)	7.9 (7.43)	4.4 (2.98)	5.6 (4.04)	7.5 (6.44)
		Case C			Case D	
1	20.2 (1.57)	9.5 (3.42)	6.4 (4.92)	28.2 (2.05)	6.0 (3.97)	6.3 (6.38)
3	28.6 (1.32)	18.0 (3.60)	8.5 (5.53)	3.8 (3.01)	5.9 (4.24)	7.4 (6.85)
5	41.3 (1.50)	14.0 (4.21)	7.9 (6.00)	4.3 (3.01)	3.6 (3.62)	6.3 (5.31)

Table 3. Power estimates of nominal 5% level (A)DF tests for the trend stationary AR(1) case with $\rho = 0.90$

	DF tests			ADF tests		
	T			T		
δ	48	96	160	48 ($\bar{\hat{k}}$)	96 ($\bar{\hat{k}}$)	160 ($\bar{\hat{k}}$)
0	**6.6**	**15.6**	**42.6**	**10.9** (0.57)	**20.2** (1.65)	**40.5** (3.22)
				case A: $\delta = \delta_1 = -\delta_2,\ \delta_3 = \delta_4 = 0$		
1	10.2	23.7	55.9	13.4 (0.58)	24.7 (2.02)	42.0 (4.69)
3	44.3	75.4	97.2	19.7 (2.94)	21.7 (6.26)	28.7 (9.84)
5	84.5	98.6	100.0	9.5 (3.83)	20.3 (6.21)	27.1 (9.39)
				case B: $\delta = \delta_1 = -\delta_3,\ \delta_2 = \delta_4 = 0$		
1	10.1	24.4	55.6	13.3 (0.56)	23.7 (2.13)	38.4 (4.87)
3	36.9	72.8	96.6	12.3 (2.52)	18.4 (5.50)	24.4 (8.68)
5	72.2	97.6	100.0	6.3 (2.83)	19.7 (5.23)	24.9 (7.89)
				case C: $\delta = \delta_1 = \delta_3 = -\delta_2 = -\delta_4$		
1	15.3	33.4	68.3	16.6 (1.05)	24.4 (3.42)	34.0 (6.55)
3	76.9	95.9	99.9	15.9 (2.16)	22.4 (4.17)	31.7 (6.49)
5	98.3	100.0	100.0	12.6 (1.94)	25.0 (3.39)	32.9 (5.34)
				case D: $\delta = \delta_1 = \delta_2 = -\delta_3 = -\delta_4$		
1	13.0	32.5	66.8	15.0 (0.80)	23.8 (3.31)	30.9 (6.91)
3	52.8	91.7	99.7	7.7 (2.90)	19.6 (5.47)	24.3 (8.30)
5	80.6	99.8	100.0	4.7 (2.72)	21.2 (4.72)	26.9 (6.90)

Table 4. Power estimates of nominal 5% level ADF tests for more realistic trend stationary processes

		T			T	
δ	48 ($\bar{\hat{k}}$)	96 ($\bar{\hat{k}}$)	160 ($\bar{\hat{k}}$)	48 ($\bar{\hat{k}}$)	96 ($\bar{\hat{k}}$)	160 ($\bar{\hat{k}}$)
DGP: $y_t = \alpha + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] D_{tj} + 0.05\, t + 0.90\, y_{t-4} + \epsilon_t$						
0	**4.5** (3.05)	**5.3** (3.73)	**16.0** (5.03)	**4.5** (3.05)	**5.3** (3.73)	**16.0** (5.03)
		Case A			Case B	
1	5.1 (3.05)	6.2 (3.99)	15.8 (5.77)	4.1 (3.06)	5.9 (3.74)	16.8 (4.95)
3	6.8 (3.21)	6.5 (5.63)	14.1 (8.17)	6.3 (3.04)	6.2 (4.66)	15.9 (7.13)
5	14.9 (3.29)	6.0 (6.79)	13.0 (9.19)	13.4 (3.05)	6.3 (5.87)	14.7 (8.53)
		Case C			Case D	
1	7.5 (3.05)	5.5 (4.92)	14.8 (7.09)	3.7 (3.06)	5.5 (3.71)	16.3 (4.98)
3	11.4 (3.48)	6.0 (5.99)	15.0 (8.12)	8.3 (3.09)	6.5 (5.67)	15.1 (8.36)
5	8.1 (3.80)	6.0 (5.74)	16.0 (7.43)	21.3 (3.23)	6.1 (6.11)	14.2 (8.51)
DGP: $y_t = \alpha + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] D_{tj} + 0.05\, t + 0.41\, y_{t-2} + 0.324\, y_{t-4} + \epsilon_t$						
0	**41.7** (1.16)	**35.0** (3.10)	**40.0** (4.87)	**41.7** (1.16)	**35.0** (3.10)	**40.0** (4.87)
		Case A			Case B	
1	39.5 (1.42)	21.2 (3.42)	35.2 (5.01)	44.7 (1.45)	20.0 (3.59)	34.7 (5.44)
3	14.0 (2.84)	17.7 (4.37)	32.9 (7.15)	10.0 (2.88)	17.7 (4.42)	33.6 (7.29)
5	9.8 (3.12)	19.0 (4.91)	36.4 (7.35)	7.1 (2.95)	15.8 (4.00)	35.8 (6.37)
		Case C			Case D	
1	31.1 (1.50)	26.6 (3.32)	38.0 (4.88)	37.9 (1.88)	16.7 (3.93)	33.6 (6.36)
3	42.8 (1.25)	41.8 (3.49)	42.9 (5.35)	5.9 (2.98)	16.4 (4.25)	35.5 (6.79)
5	57.3 (1.37)	30.5 (4.10)	36.3 (5.89)	6.7 (2.99)	11.9 (3.61)	33.5 (5.24)
DGP: $y_t = \alpha + [\sum_{j=1}^{4} \gamma_j + I(t > \tau) \sum_{j=1}^{4} \delta_j] D_{tj} + 0.05\, t + 0.90\, y_{t-4} + \epsilon_t - 0.4\epsilon_{t-4}$						
0	**14.1** (2.87)	**3.5** (5.83)	**18.9** (8.38)	**14.1** (2.87)	**3.5** (5.83)	**18.9** (8.38)
		Case A			Case B	
1	8.2 (3.02)	4.3 (5.83)	19.8 (8.48)	6.4 (3.09)	4.4 (6.39)	20.9 (8.41)
3	8.5 (3.09)	4.1 (5.26)	20.6 (9.19)	9.4 (3.08)	4.4 (4.53)	20.3 (8.53)
5	18.8 (3.12)	5.1 (5.99)	18.9 (10.39)	18.7 (3.09)	3.5 (5.12)	22.3 (9.24)
		Case C			Case D	
1	29.6 (2.30)	3.7 (5.09)	19.9 (8.77)	6.0 (3.09)	3.7 (6.06)	19.8 (7.98)
3	38.0 (2.47)	5.1 (6.46)	22.6 (9.52)	12.7 (3.15)	4.5 (4.98)	23.6 (9.16)
5	34.0 (2.93)	5.4 (6.41)	23.2 (8.86)	29.3 (3.30)	4.6 (5.23)	23.1 (9.24)