TESTING FOR INVERTIBILITY IN UNIVARIATE ARIMA PROCESSES

Rafael Flores de Frutos(*)
Miguel Jerez Méndez(*)

May 2002

Abstract. We propose a test statistic to detect whether a differenced time series follows an invertible ARIMA process. The test has a standard χ^2_1 distribution under the null, it is easy to compute and shows an excellent performance when compared with the augmented Dickey-Fuller statistic.

Keywords. Invertibility; Overdifferencing; Unit Root Tests.

JEL classification codes: C12, C22

(*) Departamento de Economía Cuantitativa. Facultad de Ciencias Económicas y Empresariales. Universidad Complutense de Madrid. Campus de Somosaguas, 28223 Madrid (SPAIN). Phone: +34 91 394 23 70. Fax: +34 91 394 26 13. Rafael Flores de Frutos e-mail: eccua04@sis.ucm.es.

1. Introduction

In determining the degree of differencing of a time series two different test-based approaches compete.

First, the standard AR approach starts from a potentially subdifferenced time series and tests the null of a unit root in the AR polynomial of the corresponding ARIMA representation. Some outstanding representatives of this approach are the LBI (Local Best Invariant) statistic, LBIU (Local Best Invariant Unbiased) statistic, Dickey-Fuller (DF hereafter) test, von Neumann ratio, Durbin-Watson test or the modified Durbin-Watson test. See Nabeya and Tanaka (1990) or Tanaka (1996, Chap. 9), for a discussion on these tests. Among them, the most widely used are the DF statistic or its augmented version. They does not have a standard distribution but are easy to compute and modify to deal with size distortions if the data generating process (DGP) is not a pure AR.

Second, the MA approach focuses in a potentially overdifferenced time series and then tests the null of a unit root in the MA polynomial of its ARIMA representation. Important works in this line are those of Arellano and Pantula (1990), Tanaka (1990), Tsay (1993) or Saikkonen and Luukkonen (1993), SL hereafter. The R1 and R2 tests proposed in SL belong to the LBI class, their power is higher than any other test into the MA approach and, when the data generating process is a univariate ARIMA model without deterministic components, they have χ^2 asymptotic distributions under the null, see Tanaka (1996, p. 376 and 384).

In this paper we explore the ability of a forecast error-based test, originally devised to detect a structural change (Lütkepohl 1993, p. 387), to detect noninvertibility situations.

The basic idea consists of comparing the forecast errors resulting from a likely overdifferenced model with those obtained from the corresponding subdifferenced model. The null is rejected if the squared standardized forecast error is too big when compared with a percentile of a χ^2 distribution. This works because overdifferencing generates unit roots in the MA polynomial, so a finite AR model cannot approximate the DGP. Therefore, applying least squares to the AR approximation yields inconsistent estimates of the parameters and poor one-step-ahead forecast errors.

The proposed test shares some nice features of both DF and SL tests: a) it has a standard distribution under the null; b) it is easy to compute; c) when the DGP is not a pure AR process it shows lower size distortions than the ADF test for any lag length choice; and d) its performance is similar to that of the powerful R1 and R2 tests proposed in SL.

The article is organized as follows. Section 2 describes the proposed test statistic. Section 3 illustrates its performance in finite samples. Finally, Section 4 presents the most important conclusions.

2. Three important results

Consider the stationary ARMA model for the nonseasonal time series z_t:

$$\varphi_p(B)\ z_t = \theta_q^z(B)\ a_t \tag{1}$$

where $z_t \equiv \nabla y_t$, $\nabla \equiv 1-B$, B is the lag operator and a_t is a white noise process with variance σ^2.

The roots of $\varphi_p(B)=0$ are assumed to lie outside the unit circle, but the polynomial $\theta_q^z(B)$ might have a factor ∇, i.e., the process is stationary but it might be noninvertible. Under these conditions, y_t follows the stationary and invertible ARMA(p, q-1) process:

$$\varphi_p(B)\, y_t = \theta_{q-1}^y(B)\, a_t$$

with

$$\theta_q^z(B) = \nabla\theta_{q-1}^y(B).$$

On the other hand if ∇ is not a factor of $\theta_q^z(B)$, y_t follows the ARIMA (p,1,q):

$$\varphi_p(B)\nabla y_t = \theta_q^z(B)\, a_t$$

Result 1: Assuming that:

$$\begin{aligned}
&a)\ L \text{ is a function } (n_T) \text{ of the sample size T}\\
&b)\ n_T \to \infty \quad as \quad T \to \infty\\
&c)\ \frac{n_T^3}{T} \to 0 \quad as \quad T \to \infty\\
&d)\ \sqrt{T} \sum_{i=n_T+1}^{\infty} |\varphi_i| \to 0 \quad as \quad T \to \infty
\end{aligned} \tag{2}$$

whatever the process followed by y_t, it can be approximated by the AR(L) process:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_L y_{t-L} + u_t \tag{3}$$

Conditions (2a)-(2d) assure that, as $T \to \infty$, u_t approaches a_t and least squares (LS) to model (3) yields consistent estimates of φ_i, $i=1,2,\dots,L$, see Lütkepohl (1993, p. 306-309).

Result 2: Only under invertibility, the process (1) can be approximated by a finite AR process:

$$z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \dots + \pi_{L-1} z_{t-(L-1)} + r_t \tag{4}$$

Note that condition (2d) holds only if the roots of $\theta_q^z(B)=0$ lie outside the unit circle. Thus, only under invertibility LS to (4) yields consistent estimates of π_i, $i=1,2,\dots,L-1$.

Result 3: Consider the quantity:

$$\tau_1 = \frac{u_{L+1}^2}{\sigma^2} \tag{5}$$

where u_{L+1} and σ^2 are the one-step forecast error at period L+1 from the AR model (3) and its variance, respectively. Under normality, τ_1 follows a standard χ_1^2 distribution, see Lütkepohl (1993, p. 387).

Based on results 1 to 3, we propose the following statistic for testing the null hypothesis of invertibility:

$$\lambda(K) = (T-2L-1)\left[\frac{\dfrac{\sum_{t=L+1}^{L+K}\hat{r}_t^2}{T-2L+1} - \dfrac{\sum_{t=L+2}^{L+K}\hat{u}_t^2}{T-2L-1}}{\dfrac{\sum_{t=L+2}^{T}\hat{u}_t^2}{T-2L-1}}\right] \tag{6}$$

where:

1) $\hat{r}_t$ are the LS residuals when (4) is estimated with *T-L* effective number of observations, *t=L+1,L+2, ..., T*. The degrees of freedom in this regression are *T-2L+1*.

2) $\hat{u}_t$ are the LS residuals when (3) is estimated with *T-L-1* effective number of observations, *t=L+2, L+3, ..., T*. The degrees of freedom in this regression are *T-2L-1*.

3) *K* is an arbitrary constant and it is <u>not</u> a function of *T*.

Under invertibility and for any choice of *K:*

$$plim\left[\frac{T-2L-1}{T-2L+1}\sum_{t=L+1}^{L+K}\hat{r}_t^2 - \sum_{t=L+2}^{L+K}\hat{u}_t^2\right] = plim\ [u_{t+L}^2] \tag{7a}$$

$$plim\left[\frac{\sum_{t=L+2}^{T}\hat{u}_t^2}{T-2L-1}\right] = \sigma^2 \tag{7b}$$

Then under invertibility, $plim[\tau_1 - \lambda(K)] = 0$, implying that $\lambda(K)$ and τ_1 have the same asymptotic χ_1^2 distribution. Note that, as we mentioned before, invertibility is necessary for LS to yield consistent estimates in (4) and thus for (7a) to hold.

In practice we will have a fixed sample size *T*. If the constant *K* is set equal to its maximum, i.e. *K=T-L*, $\lambda(K)$ will be:

$$\lambda(T-L) = (T-2L-1)\left[\frac{\dfrac{\sum_{t=L+1}^{T}\hat{r}_t^2}{T-2L+1} - \dfrac{\sum_{t=L+2}^{T}\hat{u}_t^2}{T-2L-1}}{\dfrac{\sum_{t=L+2}^{T}\hat{u}_t^2}{T-2L-1}}\right] \tag{8}$$

which is similar to the structural change test based on recursive residuals, see Harvey (1976). Note also that when setting K=T-L we are not proposing to fix K as a function of T, invalidating the proof about the distribution of $\lambda(K)$. Expression (8) is just an operative version of (6). On the other hand, the choice of L is expected to affect both, the empirical size and the power of our test. Next Section contains some simulation results on this issue.

3. Simulation exercise

In this section we study the performance of our test in finite samples. Using simulations we compare the size distortions of $\lambda(T-L)$ with: a) those of the Augmented Dickey-Fuller[1] (ADF) test for different values of L and b) the R1 and R2 test statistics proposed in SL.

Following SL, the time series y_t is generated according to:

$$
\begin{aligned}
y_1 &= u_1 \\
\Delta y_t &= u_t - \theta u_{t-1} \quad for\ t = 2, \ldots, T
\end{aligned}
$$

where u_t is a gaussian unobservable zero-mean error term with a stationary ARMA representation given by either:

$$u_t = \varepsilon_t + \beta \varepsilon_{t-1}$$

or

$$u_t + \alpha u_{t-1} = \varepsilon_t$$

where $\varepsilon_t \sim NID(0,1)$. The values assumed for θ are : .6, .8, .9, .95 and 1.0; the nuisance parameters α and β are: .0, $\pm$.5 and $\pm .8$. For each combination (θ, α) or (θ, β) we consider the sample sizes T = 100 and 200. Finally, we consider 12 lag lengths ($L = i \cdot T^{1/4}$; i = 1, 2, ..., 12) for each T. The nominal 5% significance level is used throughout and the number of replications is 1000.

Tables 1 and 2 illustrate the performance of the test for $\beta = \alpha = 0$ and T=100, 200 respectively. Tables 3 to 10 illustrate the performance of λ for each value of α. Figures in any column such that $\theta < 1$ represent empirical sizes related to a particular value of θ. Figures in the column $\theta = 1$ represent empirical powers, i.e. the probability of rejecting invertibility, being false this hypothesis. Note that this probability can be seen as the percentage of success in detecting noninvertibility (PNI, hereafter) associated to each choice of L and non invertible DGP. Tables 11 to 18 illustrate the performance of λ for each value of β. As in Tables 3 to 10, figures represent empirical sizes, except when $\theta = 1$ that represent empirical powers.

The most important results of this exercise are in Tables 1 and 2, so let us analyse them in detail. Columns 2-9 in Tables 1 and 2 contain the size distortions in both λ and ADF tests computed with different values of L (column 1), when the DGP for y_t is an IMA(1,1) process with θ ranging from .6 to .95. The last two columns contain the powers of each test when a white noise process is the DGP.

Two clear conclusions arise. First, our test shows lower size distortions than the ADF test for any choice of L. Second, while the size of λ become stable around the theoretical 5% from a particular value of L, the ADF test does not. Further, it seems to have problems for reaching this size when T=100. Therefore, ADF seems to be more sensitive to the lack of degrees of freedom than λ, experiencing important size distortions when L is too large. These features become clearer when θ is close to one. For T=200 (Table 2) previous remarks are still valid, but the size distortions in the ADF, due to the degrees of freedom effect, are now less apparent. Similar conclusions can be drawn from the remaining Tables where the DGP is either a mixed ARIMA or a pure MA model. A low power has to be accepted when a large L is needed for the AR approximation, unless T is large enough. If T is small or moderate a 5% size might not be attainable by the ADF test.

Given that: a) the larger α (or the smaller β) the larger the number of lags needed to get the 5% nominal size, b) the larger L the lower the power (or PNI in this case) and c) when T is not be very large, it will be necessary to have an *a priori* idea about the univariate DGP. In this task, the Box and Jenkins (1976) approach for ARIMA model-building can be very useful. Two versions of the DGP can be then proposed, one of them likely subdifferenced, then a standard simulation exercise can be carried out. The objective is to minimize the size distortions by choosing the appropriate L keeping the power of our test as high as possible.

Finally, Tables 19 and 20 summarize the performance of λ versus R1 and R2 proposed in SL. As λ has the opposite null hypothesis that either R1 or R2, the comparison is not easy. Figures in these Tables represent percentages of success in detecting invertibility (PDI, hereafter). In the case of R1 or R2, these figures mean power (computed for a theoretical percentage of success in detecting noninvertibility: PNI=95%) when the DGP is invertible. In the case of λ, PDI figures are computed as 100 minus the size distortion associated to the value of L that guarantee a PNI $ 95%. This can be done by choosing L small enough. Thus, taking as DGP any invertible process in Tables 1-18 (except those with θ=.6 for saving space) and having set: (a) the PNI=95% for R1 and R2 and (b) PNI $ 95% for λ, we study the success of the three tests in detecting invertibility. For instance, when the DGP for z_t is a MA(1) with θ ' .95 and T=100, the PDI (powers) of R1 and R2 are 57% and 32% respectively (figures taken from SL). For this DGP, $L=3\ T^{1/4}=9$ corresponds to the lag length for which λ shows the closest but larger PNI to 95% (95.9%). For this L, the size distortion is 42.6% (see Table 3) which implies an approximate PDI of 57%, equal to that of R1.

From these two Tables it can be concluded that the ability of λ for detecting invertibility, conditional to, at least, a 95% ability of detecting noninvertibility, is similar to that of SL tests, R1 and R2. And this seems to be the case for all DGP and sample sizes considered.

There are two questions, arising from this analysis, that are now under research: a) is there an optimal K? and b) as τ_1 is a subfamily of τ_h, see Lütkepohl (1993, p. 387), a test analogous to $\lambda(K)$ with the same asymptotic distribution as τ_h can be defined, therefore, is there any gain in choosing h>1?

4. Conclusions

In this paper we propose a test for invertibility of an ARIMA(p,1,q) process. This test has some nice features: First, it is easy to compute as it only requires a LS routine. Second, it follows a χ^2_1 asymptotic distribution if there are no deterministic components in the series. Third, when compared with the ADF test, it shows lower size distortions for any choice of L. Fourth, it always attains the 5% theoretical size, i.e. there exists a value of L from which the empirical size stabilizes around 5%. This is not the case for the ADF test when T is moderate (T=100) and the DGP has moving average roots close to 1. Finally, a comparison with R1 and R2, two of the most powerful tests into the MA approach proposed by Saikkonen and Luukkonen (1993), indicates that ours performs quite well.

References

Arellano, C. and S.G. Pantula (1990) Trend Stationarity Versus Difference Stationarity, in: Proceedings of the Business and Economic Statistics Section (American Statistical Association) 188-196.

Box, G and G. Jenkins (1976) *Time Series Analysis Forecasting and Control*, Holden Day, San Francisco.

Harvey, A.C. (1976) An Alternative Proof and Generalization of a Test for Structural Change, *The American Statistician*, 30, 122-123.

Lütkepohl, H. (1993) Introduction to Multiple Time Series Analysis , Springer-Verlag, Berlin.

Nabeya, S. and K. Tanaka (1990) Limiting Powers of Unit Root Tests in Time Series Regression, *Journal of Econometrics*, 46, 247-271.

Saikkonen, P. and R. Luukkonen (1993) Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models, Journal of the American Statistical Association, 88, 422, 596-601.

Tsay, R.S. (1993) Testing for Noninvertible Models With Applications, Journal of Business and Economic Statistics 11, 2, 225-233.

Tanaka, K. (1990) Testing for a Moving Average Unit Root, *Econometric Theory*, 6, 433-444.

Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, John Wiley & Sons, New York.

Notes

1. The ADF auxiliary regression was:

$$\Delta y_t = \rho\, y_{t-1} + \sum_{j=1}^{L-1} \gamma_j\, \Delta y_{t-j} + \varepsilon_t$$

Acknowledgements

Comments from Antonio Aznar, Halbert White and participants in the session number XVI-EC209 of the ESEM98 at Berlin are gratefully acknowledged, as well as the financial support from DGICYT PB95-0912195.

Table 1. Empirical size (and power) of λ and ADF when the DGP is $(1 \& B)y_t\ '\ (1 \& \theta B)a_t$, T=100.

	θ ' .6		θ ' .8		θ ' .9		θ ' .95		θ ' 1	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	16.3	20.1	46.5	54.3	80.1	86.4	98.6	99.3	100.0	100.0
6	7.7	9.7	17.6	23.1	40.1	49.1	74.8	84.9	99.7	100.0
9	5.3	5.5	9.4	12.2	20.3	28.4	42.6	59.0	95.9	99.0
13	6.0	5.4	6.7	8.2	9.3	15.3	23.0	36.1	60.0	85.6
16	4.7	6.1	6.3	7.5	7.1	12.1	15.2	26.3	45.6	76.1
19	5.4	6.3	4.6	7.2	5.2	8.8	10.7	19.7	30.2	62.6
22	6.1	8.3	5.5	7.1	3.3	9.4	7.5	18.0	20.7	51.7
25	5.3	7.7	5.6	6.9	5.0	9.4	4.0	13.8	15.4	44.8
28	5.4	8.8	4.3	8.6	3.5	8.6	4.7	15.4	9.3	40.5
32	4.4	11.5	3.3	8.9	2.8	9.4	4.9	12.9	7.9	31.8
35	5.2	14.3	2.6	10.9	3.4	10.9	3.3	13.6	6.3	29.1
38	5.3	17.2	3.6	15.2	3.6	13.4	3.3	14.0	5.9	27.8

Theoretical size 5%

Table 2. Empirical size (and power) of λ and ADF when the DGP is $(1 \& B)y_t\ '\ (1 \& \theta B)a_t$, T=200.

	θ ' .6		θ ' .8		θ ' .9		θ ' .95		θ ' 1	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	9.1	11.3	36.8	43.2	72.9	80.8	97.2	98.8	100.0	100.0
8	6.1	6.3	11.1	13.6	35.8	44.6	71.0	79.6	100.0	100.0
11	5.6	5.2	7.0	8.3	19.9	25.8	52.3	62.6	99.9	100.0
15	5.1	5.0	5.6	6.7	10.0	14.1	33.1	42.4	98.7	99.9
19	5.5	4.5	4.9	5.5	7.1	9.1	21.1	30.6	90.2	98.4
23	5.4	4.9	3.8	4.9	5.2	7.9	14.5	22.3	76.2	95.2
26	4.2	5.9	4.6	4.1	5.8	7.8	8.8	14.0	64.1	86.1
30	5.3	6.3	4.9	4.6	6.4	6.9	7.1	11.3	50.0	80.8
34	4.5	6.4	4.9	5.1	6.5	7.1	7.9	12.4	38.5	72.7
38	4.5	6.0	4.2	5.6	5.7	7.0	6.6	10.7	30.4	63.4
41	5.6	6.9	5.8	6.8	4.1	7.1	5.5	9.7	24.2	58.3
45	5.6	6.2	5.4	6.6	5.7	6.3	5.4	9.7	19.2	50.1

Theoretical size 5%

Table 3. Empirical size (and power) of λ and ADF when the DGP is $(1 \& .8B)(1 \& B)y_t \; ' \; (1 \& \theta B)a_t$, T=100.

	θ ' .6		θ ' .8		θ ' .9		θ ' .95		θ ' 1	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	5.4	2.1	7.5	6.3	14.8	19.9	36.9	49.3	95.0	97.4
6	6.7	4.5	7.5	6.7	9.6	14.4	25.8	36.7	68.5	88.8
9	6.8	4.3	5.4	4.8	7.5	11.5	19.5	27.8	50.5	76.9
13	6.6	6.0	6.6	5.6	6.7	8.6	13.1	20.2	35.9	58.7
16	5.5	5.8	6.4	6.3	5.7	8.4	9.7	15.7	26.3	49.3
19	5.6	5.8	5.0	6.6	6.7	8.5	6.7	13.7	19.7	41.8
22	4.4	6.6	4.7	7.6	6.0	8.2	5.4	11.8	15.0	38.9
25	4.7	8.1	5.1	8.2	5.2	7.5	5.8	10.5	14.4	33.3
28	6.0	9.4	5.2	9.0	4.3	7.5	4.3	12.3	10.8	33.2
32	5.5	12.4	4.8	11.6	4.5	11.5	3.7	13.3	7.0	31.4
35	5.1	15.3	5.1	12.5	3.5	12.7	3.9	13.9	5.8	29.4
38	6.2	20.0	4.9	17.3	4.6	14.9	4.0	15.6	6.1	31.0

Theoretical size 5%

Table 4. Empirical size (and power) of λ and ADF when the DGP is $(1 \& .8B)(1 \& B)y_t \; ' \; (1 \& \theta B)a_t$, T=200.

	θ ' .6		θ ' .8		θ ' .9		θ ' .95		θ ' 1	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	6.0	2.5	6.6	5.6	18.8	22.9	50.0	56.8	99.9	99.9
8	6.2	4.5	6.1	5.1	11.1	15.7	32.6	40.0	98.1	99.7
11	5.9	5.4	4.8	5.3	7.9	10.6	20.4	27.3	95.8	99.1
15	6.2	5.3	4.7	5.5	6.2	7.1	14.7	20.9	80.8	95.6
19	5.5	5.5	6.0	5.9	6.4	7.0	11.9	17.6	68.2	87.7
23	4.3	5.5	5.5	5.5	7.0	5.7	9.2	13.5	55.8	81.1
26	5.7	5.9	5.4	4.7	5.1	5.3	8.6	12.7	47.1	77.6
30	5.2	6.4	5.6	5.0	6.7	6.0	6.5	10.9	38.0	68.8
34	5.3	5.6	5.2	5.6	3.9	5.5	6.7	9.7	30.1	56.3
38	5.3	6.4	4.9	6.3	4.3	5.1	6.5	9.3	24.3	52.6
41	4.4	6.6	4.9	6.2	4.3	5.8	5.5	9.4	21.2	48.8
45	5.6	7.3	3.9	7.4	4.9	5.9	4.4	9.6	18.7	44.5

Theoretical size 5%

Table 5. Empirical size (and power) of λ and ADF when the DGP is $(1-.5B)(1-B)y_t = (1-\theta B)a_t$, T=100.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	7.0	7.0	21.1	28.1	57.2	65.5	84.3	92.4	100.0	100.0
6	6.5	5.3	10.0	13.8	30.1	38.0	54.3	69.7	97.2	100.0
9	6.8	4.4	7.1	8.7	15.9	24.9	30.0	44.5	81.3	94.9
13	5.2	3.6	4.2	6.3	10.3	13.5	18.0	28.4	53.5	82.6
16	4.4	4.4	5.6	6.1	7.8	10.9	10.2	20.3	37.3	69.0
19	3.9	5.1	4.6	5.8	6.0	8.6	8.2	16.1	29.0	57.5
22	5.4	7.1	4.5	8.5	4.2	8.7	7.1	13.7	22.1	49.0
25	4.9	7.2	5.0	9.0	3.7	8.2	5.5	13.2	16.0	43.7
28	4.0	8.8	3.9	9.3	3.8	9.9	4.9	13.4	10.7	35.9
32	3.7	10.7	5.3	11.7	4.4	11.0	3.7	12.2	9.4	31.7
35	4.6	11.5	4.7	12.5	3.7	10.9	3.2	12.0	6.4	32.0
38	4.5	15.8	4.1	15.7	2.5	12.5	3.1	15.2	6.1	31.3

Theoretical size 5%

Table 6. Empirical size (and power) of λ and ADF when the DGP is $(1-.5B)(1-B)y_t = (1-\theta B)a_t$, T=200.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	6.1	5.3	18.5	24.0	50.7	59.4	87.8	93.1	100.0	100.0
8	6.1	5.0	8.7	9.7	26.4	31.6	58.2	68.3	100.0	100.0
11	6.0	4.5	6.9	7.6	15.5	23.1	41.4	51.5	99.3	100.0
15	4.9	4.4	7.1	6.5	9.5	14.9	26.3	36.2	95.5	99.2
19	5.3	5.2	6.5	6.4	7.6	10.4	18.5	26.2	84.0	96.4
23	5.2	5.3	6.1	5.9	6.5	8.6	12.9	19.8	71.2	90.4
26	5.2	4.9	5.5	6.1	6.8	8.2	9.8	14.7	56.5	85.4
30	4.9	4.9	5.9	5.5	6.2	7.5	7.3	13.3	45.4	76.2
34	6.2	7.5	3.9	4.4	5.0	6.2	7.8	11.3	37.3	69.3
38	4.8	7.6	4.4	4.5	4.8	6.6	6.0	9.5	28.7	59.7
41	4.9	7.3	3.2	4.7	4.8	7.0	5.9	9.2	24.1	54.8
45	5.4	7.9	4.5	5.1	4.6	6.9	4.4	8.3	20.9	49.9

Theoretical size 5%

Table 7. Empirical size (and power) of λ and ADF when the DGP is $(1+.5B)(1-B)y_t = (1-\theta B)a_t$, T=100.

	$\theta = .6$		$\theta = .8$		$\theta = .9$		$\theta = .95$		$\theta = 1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	20.6	24.0	59.8	67.4	92.5	96.2	100.0	100.0	100.0	100.0
6	8.0	7.7	19.3	27.2	47.2	58.9	81.2	89.8	99.7	100.0
9	4.4	5.6	8.7	12.2	23.7	34.3	48.4	63.6	95.8	98.9
13	5.8	5.5	5.5	7.2	11.8	19.1	25.6	41.4	68.8	89.3
16	5.4	6.7	4.7	6.9	8.6	14.2	16.7	30.9	46.7	77.3
19	5.9	6.3	4.9	5.9	6.0	10.1	10.9	22.1	32.4	65.3
22	5.2	7.1	4.5	6.5	4.8	9.9	8.0	19.2	23.5	52.2
25	4.3	8.5	5.6	7.0	4.6	9.1	7.6	16.1	15.5	46.2
28	5.6	8.2	4.5	8.9	3.5	9.6	4.7	12.9	11.6	39.1
32	4.8	10.6	5.5	9.2	3.0	10.3	2.5	10.8	8.8	35.6
35	4.0	12.9	3.2	11.9	2.8	10.1	2.8	11.8	5.7	32.5
38	4.2	15.4	3.2	12.8	2.5	13.3	2.2	13.9	5.5	30.0

Theoretical size 5%

Table 8. Empirical size (and power) of λ and ADF when the DGP is $(1+.5B)(1-B)y_t = (1-\theta B)a_t$, T=200.

	$\theta = .6$		$\theta = .8$		$\theta = .9$		$\theta = .95$		$\theta = 1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	12.9	16.3	44.0	49.5	84.5	90.1	99.0	99.9	100.0	100.0
8	6.2	6.8	14.7	17.3	41.0	50.8	75.5	84.3	100.0	100.0
11	5.3	4.5	8.3	10.8	21.4	27.3	53.8	65.0	100.0	100.0
15	5.5	5.0	8.0	8.2	13.0	16.7	33.9	44.5	98.1	100.0
19	5.7	4.9	5.3	7.1	8.0	11.9	22.4	32.2	89.9	97.9
23	5.8	5.0	6.4	5.9	7.6	8.8	14.0	22.5	79.6	93.8
26	4.5	5.4	5.7	6.2	6.6	8.1	9.2	18.4	64.3	88.5
30	5.2	6.1	5.6	5.6	5.2	8.0	9.1	15.4	51.0	80.9
34	5.9	5.6	5.7	5.5	5.0	7.0	5.4	10.6	43.0	74.6
38	4.3	5.9	5.4	5.7	4.7	6.1	5.1	10.4	33.5	66.3
41	5.4	5.6	4.7	5.3	4.1	6.6	4.9	9.7	27.4	59.6
45	4.8	5.6	5.8	5.8	3.8	5.9	4.6	8.1	19.8	52.4

Theoretical size 5%

Table 9. Empirical size (and power) of λ and ADF when the DGP is $(1\,\%\,.8B)(1\,\&\,B)y_t\ '\ (1\,\&\,\theta B)a_t$, T=100.

	θ ' .6		θ ' .8		θ ' .9		θ ' .95		θ ' 1	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	21.6	27.6	64.2	71.4	93.9	97.5	99.9	100.0	100.0	100.0
6	7.3	7.5	19.3	27.4	49.1	60.5	82.7	91.8	100.0	100.0
9	6.6	6.0	10.0	12.7	23.7	34.3	49.4	64.2	95.3	99.4
13	6.0	5.3	6.6	7.8	10.0	17.1	23.7	39.4	67.0	91.2
16	6.3	6.2	5.0	7.2	6.3	12.0	13.4	25.1	48.9	79.3
19	6.3	7.5	5.6	7.1	4.8	10.3	9.0	19.6	33.4	64.7
22	4.9	6.9	5.7	7.7	4.0	8.8	7.2	17.9	23.3	54.6
25	4.6	7.1	4.3	8.5	4.2	8.3	6.3	15.3	15.3	46.2
28	4.7	8.8	4.3	8.6	3.7	8.2	3.6	11.2	12.7	42.3
32	3.2	10.4	3.5	9.6	2.6	8.1	3.1	11.5	6.7	33.8
35	4.6	11.9	3.2	10.9	2.5	10.0	2.7	10.7	5.0	31.3
38	4.6	15.4	4.4	13.1	3.8	11.7	3.1	13.7	4.5	30.0

Theoretical size 5%

Table 10. Empirical size (and power) of λ and ADF when the DGP is $(1\,\%\,.8B)(1\,\&\,B)y_t\ '\ (1\,\&\,\theta B)a_t$, T=200.

	θ ' .6		θ ' .8		θ ' .9		θ ' .95		θ ' 1	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	14.2	17.6	50.5	58.1	84.9	90.1	99.6	99.9	100.0	100.0
8	5.9	5.3	14.2	18.1	40.2	48.2	78.4	86.2	100.0	100.0
11	5.9	5.4	9.2	9.2	20.9	28.1	56.3	65.9	99.9	100.0
15	4.6	5.4	6.6	6.2	12.2	14.7	33.0	43.9	98.6	99.9
19	5.3	4.8	4.5	4.9	7.8	10.8	20.6	30.1	90.9	98.3
23	5.1	4.9	4.5	5.1	7.4	8.4	12.5	21.5	76.4	94.5
26	5.0	5.2	5.2	6.1	7.1	9.7	10.2	17.0	65.4	90.4
30	5.7	4.7	6.5	5.7	6.4	7.7	8.0	13.3	52.9	80.7
34	5.6	7.3	5.9	7.0	6.7	7.0	7.4	11.2	38.1	74.1
38	5.6	6.3	5.4	7.6	5.1	6.6	5.5	9.3	31.6	65.7
41	4.2	6.8	5.2	7.2	4.8	7.1	4.2	8.7	27.0	58.3
45	4.6	7.7	4.0	7.6	4.7	7.1	4.0	7.4	22.6	53.3

Theoretical size 5%

Table 11. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1-.8B)(1-\theta B)a_t$, T=100.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	89.5	94.5	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0
6	35.6	45.7	75.5	82.8	97.6	99.0	100.0	100.0	100.0	100.0
9	14.7	20.7	35.6	45.4	70.7	82.8	94.8	98.3	100.0	100.0
13	7.0	9.1	12.0	20.8	32.0	45.3	59.5	77.0	96.4	100.0
16	4.9	8.3	7.1	11.9	16.3	29.2	35.0	55.8	80.2	95.8
19	4.6	7.2	5.2	8.6	9.0	18.4	20.6	37.9	56.3	87.8
22	4.4	5.4	3.0	7.4	6.6	13.6	12.5	29.8	37.0	76.5
25	4.1	5.1	3.7	6.3	4.2	10.7	7.5	21.5	24.5	62.8
28	4.2	7.3	2.9	6.4	4.5	10.6	5.1	18.2	17.3	54.8
32	4.2	10.0	4.2	6.8	4.1	9.6	3.4	15.4	9.8	43.3
35	4.1	11.0	2.9	9.5	3.0	10.2	3.3	15.1	6.4	36.6
38	3.3	12.4	1.9	10.2	2.2	11.6	3.4	15.2	5.5	35.3

Theoretical size 5%

Table 12. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1-.8B)(1-\theta B)a_t$, T=200.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	75.9	81.7	97.8	99.4	100.0	100.0	100.0	100.0	100.0	100.0
8	23.1	29.1	55.1	63.6	90.0	94.1	99.8	99.9	100.0	100.0
11	10.8	13.8	28.4	34.5	64.2	71.9	92.8	97.1	100.0	100.0
15	7.0	7.5	11.8	15.6	32.3	42.9	71.0	80.0	100.0	100.0
19	5.3	6.4	6.6	9.0	15.8	22.5	44.9	57.0	99.6	99.9
23	5.9	5.8	5.8	6.8	9.3	14.4	27.6	41.4	96.2	99.8
26	3.3	4.8	6.1	6.7	6.4	10.2	17.1	28.4	87.9	98.4
30	4.2	5.3	7.3	6.0	5.4	8.2	13.2	21.3	73.0	94.4
34	5.6	5.5	4.9	4.7	5.2	7.8	10.2	15.9	53.3	85.6
38	5.7	6.5	4.3	4.8	4.6	7.5	7.3	11.7	43.1	74.8
41	4.3	6.3	4.3	4.8	4.1	6.4	5.9	11.0	34.0	70.3
45	4.6	7.6	4.9	5.7	5.4	6.5	5.5	9.6	25.7	61.2

Theoretical size 5%

Table 13. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1-.5B)(1-\theta B)a_t$, T=100.

	$\theta = .6$		$\theta = .8$		$\theta = .9$		$\theta = .95$		$\theta = 1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	41.4	48.8	77.2	83.5	98.9	99.6	100.0	100.0	100.0	100.0
6	8.7	11.3	26.3	34.7	64.4	75.8	91.1	95.6	100.0	100.0
9	5.9	7.7	11.0	16.4	33.9	45.2	59.6	73.2	97.8	99.8
13	5.6	5.2	5.6	7.4	15.0	24.0	29.5	43.4	74.6	92.5
16	3.9	5.0	4.3	6.6	9.2	16.0	18.9	32.2	53.3	84.0
19	3.9	6.4	4.1	5.8	6.5	12.6	11.5	24.6	36.0	70.6
22	4.5	6.2	5.1	6.4	4.3	10.0	5.3	17.3	26.2	58.8
25	3.9	6.8	4.0	6.5	4.1	9.1	4.9	13.1	17.8	48.1
28	6.0	8.5	4.5	9.6	4.7	8.2	4.7	13.4	11.0	40.7
32	5.3	11.8	4.0	10.7	3.8	8.7	3.3	11.4	8.1	37.8
35	3.8	13.0	3.7	12.4	3.6	9.5	3.0	11.5	7.4	31.4
38	3.4	15.8	3.6	15.0	4.1	11.9	2.5	13.1	4.6	29.9

Theoretical size 5%

Table 14. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1-.5B)(1-\theta B)a_t$, T=200.

	$\theta = .6$		$\theta = .8$		$\theta = .9$		$\theta = .95$		$\theta = 1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	21.4	29.0	63.6	70.7	94.6	97.1	100.0	100.0	100.0	100.0
8	7.0	7.4	15.4	21.9	53.4	62.4	84.8	90.8	100.0	100.0
11	5.9	4.9	10.2	11.9	26.9	34.8	63.8	74.2	100.0	100.0
15	4.1	4.5	6.3	7.6	12.5	17.6	38.4	51.0	99.5	99.9
19	4.6	4.8	5.3	5.9	7.8	10.8	23.4	34.3	95.3	99.4
23	3.9	4.3	4.9	6.2	5.0	8.6	13.9	24.6	81.6	96.6
26	4.5	4.9	4.7	4.6	6.4	8.0	12.0	19.2	69.7	91.3
30	4.4	4.7	5.1	5.3	4.5	6.8	10.0	14.5	56.1	81.9
34	5.3	6.3	5.9	5.1	5.1	5.3	7.2	12.8	45.1	75.2
38	6.3	6.2	3.9	5.9	3.5	6.4	6.4	10.4	33.5	66.7
41	4.5	5.7	4.2	6.3	4.2	6.6	5.6	9.9	26.6	58.7
45	5.1	5.4	4.8	6.8	4.8	7.3	5.5	11.1	21.1	51.3

Theoretical size 5%

Table 15. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1+.5B)(1-\theta B)a_t$, T=100.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	7.1	8.3	35.5	43.8	70.2	79.8	94.2	97.6	100.0	100.0
6	5.7	5.1	14.9	20.9	35.4	46.2	68.6	79.8	99.3	99.9
9	6.8	4.9	8.5	11.3	18.9	27.0	40.1	55.1	94.6	97.4
13	5.7	4.7	6.9	7.9	10.2	17.0	18.4	32.7	60.0	84.1
16	6.1	5.5	6.9	6.5	6.3	14.0	13.5	21.6	46.2	72.7
19	5.7	5.0	5.2	6.9	5.9	11.2	9.2	17.8	31.8	59.7
22	3.4	6.1	5.0	8.2	4.7	8.5	7.2	14.8	20.7	50.7
25	4.5	7.0	5.8	8.6	3.7	7.9	6.2	13.0	13.5	42.6
28	4.4	8.9	3.8	9.1	4.0	7.3	4.3	11.5	11.9	38.5
32	4.7	12.2	4.7	10.4	4.4	9.6	3.7	11.3	8.8	33.5
35	3.2	14.1	3.0	12.3	3.0	10.8	3.0	12.5	5.9	32.1
38	3.9	16.9	3.1	14.8	4.1	15.2	2.8	12.9	4.8	31.5

Theoretical size 5%

Table 16. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1-.5B)(1-\theta B)a_t$, T=200.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	8.5	10.6	26.7	33.9	68.0	75.6	94.8	97.6	100.0	100.0
8	4.9	5.7	9.5	12.3	33.5	40.0	66.1	75.3	100.0	100.0
11	6.9	6.8	6.6	8.1	18.2	23.2	48.4	56.3	100.0	100.0
15	6.7	6.4	5.3	5.7	10.8	14.7	30.1	40.0	96.9	99.7
19	6.7	6.4	4.4	5.0	7.9	10.8	17.4	28.6	88.2	98.2
23	7.2	7.0	4.5	4.4	6.1	8.7	14.6	20.4	75.8	93.8
26	4.4	6.2	6.1	5.1	7.4	7.4	10.7	16.7	62.1	87.0
30	4.7	6.3	5.8	5.8	5.9	7.6	8.5	12.9	49.5	78.0
34	5.4	5.1	3.8	5.4	5.2	7.3	6.2	10.3	42.4	72.9
38	5.8	4.9	4.0	5.8	5.3	7.8	4.5	8.9	32.2	63.5
41	5.7	5.4	5.2	5.7	5.7	8.0	5.6	8.4	27.4	58.9
45	3.8	6.1	5.2	6.5	4.4	6.7	4.4	7.8	22.1	51.2

Theoretical size 5%

Table 17. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1+.8B)(1-\theta B)a_t$, T=100.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
3	7	6.8	23.7	30.3	63.3	72.2	89.8	95.2	100	100
6	7.6	7.2	14.2	17.6	37.0	49.2	69.8	81.3	99.0	100.0
9	5.2	5.2	7.6	10.4	16.6	23.7	35.1	49.8	95.5	97.3
13	5.7	5.6	6.6	8.3	8.4	14	19.8	32.4	54.6	81.9
16	4.6	6	6.1	7.6	6.8	11.3	13.9	25.6	45.1	72.8
19	5.5	5.5	4.8	7.1	5	8.4	10	19.2	27.5	58.9
22	4.3	6.5	4	6.9	4.8	8	6.4	15.4	20.6	50.3
25	4.5	7.1	2.7	6.6	4.6	8	4.9	13.2	13.9	42.1
28	3.7	8.8	3.5	8.2	4.6	9.6	4.7	12	12.3	41.5
32	5.5	11.4	4.1	9.4	3.8	11	3.8	12.2	8.3	36.7
35	4.3	15.1	3.4	11.3	4.6	12.8	2.7	11.6	6.2	33.7
38	4.7	18.1	3.2	14.2	3.3	13.6	3.1	14.4	6.3	31.6

Theoretical size 5%

Table 18. Empirical size (and power) of λ and ADF when the DGP is $(1-B)y_t = (1+.8B)(1-\theta B)a_t$, T=200.

	$\theta=.6$		$\theta=.8$		$\theta=.9$		$\theta=.95$		$\theta=1$	
L=	λ	ADF	λ	ADF	λ	ADF	λ	ADF	λ	ADF
4	10.5	12.9	32.5	38.3	68.7	75.4	95.7	98.2	100.0	100.0
8	6.6	6.7	11.1	14.4	30.3	38.7	65.3	73.6	100.0	100.0
11	5.0	4.2	7.9	7.6	15.4	21.6	44.4	55.9	99.8	100.0
15	4.3	4.3	5.4	5.5	7.7	10.9	27.0	38.7	96.7	99.7
19	4.6	4.5	4.7	5.7	5.2	8.6	18.1	25.6	85.9	97.5
23	4.2	4.7	5.0	5.4	5.1	6.9	11.6	20.4	74.9	92.1
26	5.2	7.0	4.9	4.9	5.8	6.4	9.1	15.1	60.8	87.5
30	6.3	6.2	4.6	4.5	4.4	6.0	7.3	11.1	50.8	79.1
34	5.1	5.2	4.7	5.9	5.7	6.3	6.9	10.2	37.9	67.1
38	5.4	5.5	3.7	6.6	6.3	5.8	6.7	10.3	31.0	60.7
41	5.9	6.0	5.0	6.0	5.2	6.7	6.0	9.3	25.1	53.8
45	5.2	6.7	4.6	6.5	5.4	6.9	5.1	9.6	20.4	49.9

Theoretical size 5%

Table 19: Percentage of success in detecting invertibility†

$\alpha =$		-8	-5	0	5	8
$\theta =$	Test	*T*=100				
95	λ‡	63 (1)	47 (2)	57 (3)	52 (3)	51 (3)
	R1	62	61	57	59	60
	R2	43	38	32	32	32
90	λ	85 (1)	70 (2)	80 (3)	76 (3)	76 (3)
	R1	72	77	73	76	75
	R2	61	62	56	60	60
80	λ	92 (1)	90 (2)	91 (3)	91 (3)	90 (3)
	R1	76	84	83	84	83
	R2	78	82	81	86	86
$\theta =$	Test	*T*=200				
95	λ‡	80 (3)	74 (4)	68 (4)	66 (4)	67 (4)
	R1	77	76	74	77	76
	R2	62	61	58	59	59
90	λ	92 (3)	91 (4)	90 (4)	87 (4)	88 (4)
	R1	84	86	85	86	86
	R2	80	85	82	84	85
80	λ	95 (3)	93 (4)	94 (4)	92 (4)	93 (4)
	R1	85	91	90	92	91
	R2	88	95	96	98	98

† The stochastic process is $z_t = u_t - \theta u_{t-1}$; $u_t + \alpha u_{t-1} = \varepsilon_t$.

‡ The number of lags (L) is $(T)^{1/4}$ times the figure in parentheses. It has been chosen so that the probability of detecting $\theta = 1$ is at least 95%.

Table 20: Percentage of success in detecting invertibility†

	$\beta =$	-8	-5	0	5	8
$\theta =$	Test	T=100				
95	λ‡	41 (4)	40 (3)	57 (3)	60 (3)	65 (3)
	R1	51	54	57	59	59
	R2	13	28	32	34	34
90	λ	68 (4)	66 (3)	80 (3)	81 (3)	83 (3)
	R1	67	70	73	75	76
	R2	36	50	56	60	62
80	λ	88 (4)	89 (3)	91 (3)	91 (3)	92 (3)
	R1	74	78	83	86	86
	R2	53	70	81	85	86
$\theta =$	Test	T=200				
95	λ‡	72 (6)	77 (5)	68 (4)	70 (4)	73 (4)
	R1	72	74	74	75	75
	R2	43	53	58	59	59
90	λ	91 (6)	92 (5)	90 (4)	89 (4)	92 (4)
	R1	82	82	85	86	87
	R2	64	76	82	84	85
80	λ	94 (6)	95 (5)	94 (4)	95 (4)	95 (4)
	R1	87	87	90	91	91
	R2	79	91	96	97	97

† The stochastic process is $z_t = u_t - \theta u_{t-1}$; $u_t = \varepsilon_{t-1} + \beta \varepsilon_{t-1}$

‡ The number of lags (L) is $(T)^{1/4}$ times the figure in parentheses. It has been chosen so that the probability of detecting $\theta = 1$ is at least 95%.