econometrics

MDPI

Article

Unit Root Tests: The Role of the Univariate Models Implied by Multivariate Time Series

Nunzio Cappuccio [1] and Diego Lubian [2,*]

1 Department of Economics and Management "Marco Fanno", via del Santo 33, Padova 35123, Italy; nunzio.cappuccio@unipd.it
2 Department of Economics, University of Verona, via dell'Artigliere 19, Verona 37129, Italy
* Correspondence: diego.lubian@univr.it; Tel.: +39-045-8028419

Academic Editor: Kerry Patterson
Received: 17 September 2015; Accepted: 16 March 2016; Published: 7 April 2016

Abstract: In cointegration analysis, it is customary to test the hypothesis of unit roots separately for each single time series. In this note, we point out that this procedure may imply large size distortion of the unit root tests if the DGP is a VAR. It is well-known that univariate models implied by a VAR data generating process necessarily have a finite order MA component. This feature may explain why an MA component has often been found in univariate ARIMA models for economic time series. Thereby, it has important implications for unit root tests in univariate settings given the well-known size distortion of popular unit root test in the presence of a large negative coefficient in the MA component. In a small simulation experiment, considering several popular unit root tests and the ADF sieve bootstrap unit tests, we find that, besides the well known size distortion effect, there can be substantial differences in size distortion according to which univariate time series is tested for the presence of a unit root.

Keywords: unit root tests; multivariate time series; cointegration

JEL: C32, C22, C12, C52

1. Introduction

It is well known that unit root tests may have large size distortion when the autoregressive parameter is close to unity and/or when there is a large MA component (see, for instance, [1]). The simulation evidence on the size distortion of some standard univariate unit root tests, such as the ADF test and Phillips-Perron Z_α and Z_t, is overwhelming (see, among others, [2]). Most simulation studies consider univariate unit root processes, but the same findings on the size distortion of unit root tests have been obtained by Reed [3,4] who considers a cointegrated VAR. He finds that the size distortion can be very large and not necessarily of similar magnitude across tests and univariate time series derived from the VAR. For instance, in the bivariate case, unit root tests applied to the one component may have an effective size as large as 90% while the same unit root test applied to the other component may have an effective size close to the nominal one. These finite sample results hold for DGPs characterized by quite different parameter values and for a wide range of roots of the AR component, and even in the case where the roots are 1 and 0.

In this note, we (a) provide a theoretical motivation for the finite sample size distortion observed in the presence of a large negative MA root; (b) give additional simulation evidence on its extent comparing standard and bootstrap unit root tests and (c) provide some suggestions for empirical researchers working with univariate time series implied by a VAR data generating process.

2. The Model

We start the discussion from the VAR(1) model

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix} \tag{1}$$

or, in compact notation,

$$\boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{z}_{t-1} + \boldsymbol{u}_t,$$

where $\boldsymbol{u}_t \sim$ i.i.d.$(\boldsymbol{0}, \boldsymbol{V})$. The representation for the univariate components, the so-called "final equations" in [5], can be obtained following [6] . Considering the lag polynomial matrix $\boldsymbol{A}(L) = \boldsymbol{I} - \boldsymbol{A}L$, its determinant

$$|\boldsymbol{A}(L)| = [(1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2]$$

and the adjoint matrix

$$\boldsymbol{A}^*(L) = \begin{pmatrix} 1 - a_{22}L & a_{12}L \\ a_{21}L & 1 - a_{11}L \end{pmatrix},$$

the "final equations" for the VAR(1) model are given by

$$[(1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2] \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} 1 - a_{22}L & a_{12}L \\ a_{21} & 1 - a_{11}L \end{pmatrix} \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix}. \tag{2}$$

It follows that the univariate processes evolve as an ARMA(2,1) model with a common AR component and two distinct MA components. [1]

The magnitude and sign of the roots of the characteristic equation $|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0$ determine both the stationarity or nonstationarity of the univariate time series y_t and x_t and the existence of a cointegrating relationship between them. A necessary condition for cointegration is that the roots of the characteristic equation satisfy $\lambda_1 = 1$ and $|\lambda_2| < 1$. From this unit root constraint, we obtain the restriction

$$a_{11} = 1 - \frac{a_{12}a_{21}}{1 - a_{22}}, \tag{3}$$

which can be used to obtain the VECM representation

$$\begin{aligned} \Delta y_t &= \alpha_1(y_{t-1} - \beta x_{t-1}) + u_{t1}, \\ \Delta x_t &= \alpha_2(y_{t-1} - \beta x_{t-1}) + u_{t1}, \end{aligned}$$

where $\alpha_1 = -a_{12}a_{21}/(1 - a_{22})$, $\alpha_2 = a_{21}$ and $\beta = (1 - a_{22})/a_{21}$.

The second restriction $|\lambda_2| < 1$, *i.e.*,

$$-1 < \lambda_2 = a_{22} - \frac{a_{12}a_{21}}{1 - a_{22}} < 1 \tag{4}$$

guarantees the stationarity of the error correction mechanism. In fact, it is easy to show that

$$(y_t - \beta x_t) = \lambda_2(y_{t-1} - \beta x_{t-1}) + (u_{t1} - \beta u_{t2}).$$

---

1 In general, considering a k-dimensional VAR(p) process the univariate models will be at most ARMA$(kp, (k-1)p)$, all univariate processes share the same AR component, and an MA component is present in each univariate model. See [5] for a general treatment of this issue.

Imposing the constraints $\lambda_1 = 1$ and $|\lambda_2| < 1$, the "final equations", *i.e.*, the univariate models become

$$(1-\lambda_2 L)\begin{pmatrix}\Delta y_t \\ \Delta x_t\end{pmatrix} = \begin{pmatrix}1-a_{22}L & a_{12}L \\ a_{21}L & 1-a_{11}L\end{pmatrix}\begin{pmatrix}u_{t1} \\ u_{t2}\end{pmatrix}. \quad (5)$$

Thus, if the DGP is a bivariate VAR(1) with one unit root and one cointegrating relationship, the marginal processes for the level processes follow an ARIMA(1,1,1) model and the marginal processes for the first-difference stationary processes are ARMA(1,1) processes. [2] It follows that the autocorrelation structure of the implied marginal processes, induced by the interaction of the AR and MA roots, is deemed to affect the finite sample size and power properties of unit root test in any simulation study where the DGP is a multivariate one.

Considering the right-hand side of (5), we see how the aggregate error term for each component of $\boldsymbol{z}_t$ is the sum of an MA(1) process and a lagged white noise process

$$\begin{aligned}(1-\lambda_2 L)\Delta y_t &= \xi_{t1} = u_{t1} - a_{22}u_{t-1,1} + a_{12}u_{t-1,2},\\ (1-\lambda_2 L)\Delta x_t &= \xi_{t2} = u_{t2} - a_{11}u_{t-1,2} + a_{21}u_{t-1,1}.\end{aligned}$$

It is easy to show that both aggregate error terms on the right-hand side have the autocorrelation function of an MA(1) process so that we can write

$$\begin{aligned}\xi_{t1} &= v_{t1} + \theta_1 v_{t-1,1},\\ \xi_{t2} &= v_{t2} + \theta_2 v_{t-1,2},\end{aligned} \quad (6)$$

where v_{t1} and v_{t2} are white noise processes. By setting the first-order autocorrelation coefficient of each marginal process on the left-hand side of (6) equal to the first-order autocorrelation coefficient of $\theta_i(L)v_{ti}$, *i.e.*,

$$\frac{\mathrm{Cov}(\xi_{ti}, \xi_{t-1,i})}{\mathrm{Var}(\xi_{t,i})} = \frac{\theta_i}{(1+\theta_i^2)},$$

we can find the moving average coefficients of the polynomials $\theta_i(L)$ by choosing the invertible solution of the previous second degree equation. [3]

For example, let us consider the DGP 4 in Table 1, *i.e.*,

$$\boldsymbol{A} = \begin{pmatrix}0.5 & 0.05 \\ 1 & 0.9\end{pmatrix}, \quad \boldsymbol{V} = \begin{pmatrix}0.045 & 0.017 \\ 0.017 & 0.045\end{pmatrix}.$$

The roots of the characteristic equation of the reduced form VAR are given by 1 and -0.4 and the marginal processes are given by

$$\begin{aligned}(1-0.4L)\Delta y_t &= u_{t1} - 0.9u_{t-1,1} + 0.05u_{t-1,2},\\ (1-0.4L)\Delta x_t &= u_{t2} - 0.5u_{t-1,2} + u_{t-1,1},\end{aligned}$$

which, after some algebra, can be written as

$$\begin{aligned}(1-0.4L)\Delta y_t &= v_t - 0.907v_{t-1},\\ (1-0.4L)\Delta x_t &= v_t - 0.385v_{t-1}.\end{aligned}$$

---

2 See Cubadda *et al.* [7] for a general result on the implied univariate models from cointegrated VAR and Cubadda and Triacca [8] for the $I(2)$ case.

3 A general solution to this problem for an MA process of order q has been provided in Maravall and Mathis [9].

Thus, the MA component of the process Δy_t has a large negative root, which explains the enormous size distortion, as reported in many simulation studies since Schwert [1].

3. A Simulation Experiment

In a small simulation study, we assess the size distortion of a number of unit root tests when the data comes from a cointegrated DGP. We consider the classical ADF test by Said and Dickey [10], the Z_α and Z_t tests proposed by Phillips and Perron [11], the modified MZ_α and MZ_t by Stock [12] and Perron and Ng [2], the modified Sargan-Bhargava MSB test proposed by Stock [12], the point optimal test P_T of Elliott *et al*. [13] and its modification MP_T proposed by Ng and Perron [14], and, finally the DF-GLS test by Elliott *et al*. [13]. We always estimate the spectral density at frequency zero of the error term using the autoregressive spectral density estimator as in Perron and Ng [2] and for the Z_α, Z_t, MZ_α, MZ_t, MSB tests we consider both OLS detrending and GLS detrending, as in Ng and Perron [14].

For the selection of the lag length, we do not follow a rule based just on the sample size but consider the Modified Akaike Information Criterion (MAIC) developed by Ng and Perron [14], where an upper bound to the lag length is set to the integer part of $12[(T+1)/100]^{1/4}$. Given the better performance of this information criterion compared to the BIC one, we do not consider the latter in our simulations. Further, we consider the suggestion by Perron and Qu [15] and present results for MZ_α^{PQ} obtained using OLS detrending instead of GLS detrending in the MAIC. They show that this simple modification produces tests with effective size closer to the nominal size.

We also consider two bootstrap unit-root tests. Palm *et al*. [16] carried out extensive simulation experiments on the size and power of bootstrap unit root considering ADF sieve and block bootstrap tests, based on first difference series or on residuals. Their findings suggest that, both in terms of size and power, the ADF sieve test as in Chang and Park [17] or its residual-based version perform best. In the following, we shall consider these two versions of the ADF sieve unit root test and implement the tests following the procedure set forth in Palm *et al*. [16].

Finally, to be able to make a comparison with a widely used test for the presence of unit roots and cointegration in multivariate time series, we also consider Johansen's trace-statistics, say J_{tr}, where, under the null, we should be able to reject no cointegration and not to reject the presence of one cointegrating vector.

In the simulation experiment, we consider two different ways of formulating the DGP. Firstly, as in Reed [3], we consider the VAR model in (1) subject to the constraints (3)–(4) which guarantee cointegration between y_t and x_t. Secondly, we also consider a DGP widely used in cointegration analysis (see, among others, [18]):

$$\begin{aligned} y_t - \beta x_t &= z_t, \qquad z_t = \rho z_{t-1} + e_{t1}, \\ c_1 y_t - c_2 x_t &= w_t, \quad\; w_t = w_{t-1} + e_{t2}, \\ \begin{pmatrix} e_{t1} \\ e_{t2} \end{pmatrix} &\sim \text{i.i.d.} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & \eta\sigma \\ \eta\sigma & \sigma^2 \end{pmatrix} \right) . \end{aligned} \tag{7}$$

From this parameterization, we can obtain the implied VAR(1) as

$$\begin{aligned} \begin{pmatrix} y_t \\ x_t \end{pmatrix} &= \frac{1}{\beta_1 c_1 - c_2} \begin{pmatrix} \beta c_1 - c_2\rho & -c_2\beta(1-\rho) \\ c_1(1-\rho) & c_1\beta\rho - c_2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \frac{1}{\beta_1 c_1 - c_2} \begin{pmatrix} -c_2 & \beta \\ -c_1 & 1 \end{pmatrix} \begin{pmatrix} e_{t1} \\ e_{t2} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix}, \end{aligned}$$

where the unit root constraint $a_{11} = 1 - a_{12}a_{21}/(1 - a_{22})$ is satisfied and under the condition

$$|\rho| = \left| a_{22} - \frac{a_{12}a_{21}}{(1 - a_{22})} \right| < 1,$$

we have cointegration. The VECM representation is, then, given by

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = \frac{1 - \rho}{\beta_1 c_1 - c_2} \begin{pmatrix} c_2 \\ c_1 \end{pmatrix} (y_{t-1} - \beta x_{t-1}) + \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} (y_{t-1} - \beta x_{t-1}) + \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix}.$$

From the above VAR and VECM representations, we can obtain the DGPs considered in Reed [3] in terms of the parameters of (7) and *vice versa*.

In the simulation study, we consider DGPs parameterized, as in (7), following Gonzalo [18], and as in (1), following Reed [3]. The two set of parameters are defined as follows:

- **MC(a)**: the values taken by $\boldsymbol{A}$ and $\boldsymbol{V}$ in Reed [3], reported in Table 1 together with the implied values of ρ, β, the roots of the MA component in the univariate representations and the unconditional contemporaneous autocorrelation;
- **MC(b)**: as in Gonzalo [18], we set $c_2 = -1$, $\beta = 1$ and $c_1 = 1$ and consider the following values for the remaining parameters: $\rho = (0.9, 0.5)$, $\sigma = (0.25, 1, 2)$ and $\eta = (-0.5, 0, 0.5)$, for a total of 18 experiments. The root of the common autoregressive component (one root is always equal to 1) and the coefficients of the distinct MA components of the univariate models implied by the multivariate DGP are reported in Table 2.

All results are based on a sample size $T = 100$, on 1000 simulations of the DGP and, for the bootstrap tests, on 500 bootstrap replications as in Psaradakis [19].

We notice that, for many values of the parameters, the univariate models are characterized by a large negative MA coefficient, which is exactly the circumstance in which unit root tests have low power and great size distortion even in moderately large sample sizes. For the set of parameters in **MC(b)**, this always occurs when $\rho = 0.9$ and when $\rho = 0.5$ and $\eta = -0.5$, while for the set of parameters in **MC(a)**, this occurs only in half of the cases.

Table 1. Parameter values in **MC(a)**.

	β	ρ	θ_y	θ_x	$r_{\Delta y_t, \Delta x_t}$
DGP1					
$\boldsymbol{A} = \begin{pmatrix} 0.1 & -0.1 \\ -0.9 & 0.9 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 0.061 & 0.086 \\ 0.0.86 & 0.32 \end{pmatrix}$	0.11	0	−0.71	0.12	0.258
DGP2					
$\boldsymbol{A} = \begin{pmatrix} 0.5 & 0.1 \\ 0.5 & 0.9 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 1.47 & 1.41 \\ 1.41 & 2.32 \end{pmatrix}$	−0.2	0.4	−0.92	−0.69	0.821
DGP3					
$\boldsymbol{A} = \begin{pmatrix} 0.7 & -0.9 \\ -0.3 & 0.1 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 0.045 & 0.017 \\ 0.017 & 0.045 \end{pmatrix}$	3	−0.2	0.14	−0.52	−0.195
DGP4					
$\boldsymbol{A} = \begin{pmatrix} 0.5 & 0.05 \\ 1 & 0.9 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 0.045 & 0.017 \\ 0.017 & 0.045 \end{pmatrix}$	−0.1	0.4	−0.90	−0.38	0.592

Note: $\boldsymbol{A}$ and $\boldsymbol{V}$ are defined in (1).

Some additional remarks are in order. First of all, **MC(a)** is able to generate, at least for the parameter values considered here, greater heterogeneity in the MA roots of the univariate first-difference processes than that generated by **MC(b)**. In fact, in **MC(a)**, we observe a large negative MA root for Δy_t associated to a small or a medium size negative MA root for Δx_t or a positive MA root for Δy_t associated to a negative one for Δx_t. On the contrary, in **MC(b)**, looking, for instance, at the upper panel of Table 2 (the case in which $\rho = 0.9$), we may see how the coefficients of the

MA components are almost identical for the two processes Δy_t and Δx_t and that, to a lower extent, the same applies to the lower panel ($\rho = 0.5$). When this occurs, the univariate representations of unit root processes resemble each other, and this may explain the similar behavior of unit root tests applied to y_t and x_t, separately.

Table 2. Parameter values in **MC(b)**.

	AR root $\rho = 0.9$								
		$\sigma = 0.25$			$\sigma = 1$			$\sigma = 2$	
η	θ_y	θ_x	$r_{\Delta y_t,\Delta x_t}$	θ_y	θ_x	$r_{\Delta y_t,\Delta x_t}$	θ_y	θ_x	$r_{\Delta y_t,\Delta x_t}$
−0.5	−0.972	−0.978	0.911	−0.902	−0.942	0.029	−0.885	−0.924	−0.635
0	−0.975	−0.975	0.887	−0.930	−0.930	0.025	−0.910	−0.910	−0.583
0.5	−0.978	−0.972	0.911	−0.942	−0.902	0.029	−0.924	−0.885	−0.635
	AR root $\rho = 0.5$								
		$\sigma = 0.25$			$\sigma = 1$			$\sigma = 2$	
η	θ_y	θ_x	$r_{\Delta y_t,\Delta x_t}$	θ_y	θ_x	$r_{\Delta y_t,\Delta x_t}$	θ_y	θ_x	$r_{\Delta y_t,\Delta x_t}$
−0.5	−0.872	−0.893	0.925	−0.884	−0.884	0.910	−0.893	−0.872	0.925
0	−0.565	−0.717	0.158	−0.666	−0.666	0.142	−0.717	−0.565	0.158
0.5	−0.451	−0.626	−0.539	−0.565	−0.565	−0.500	−0.626	−0.451	−0.539

Furthermore, when $\rho = 0.9$, the coefficients of the MA component in Δy_t and Δx_t are not only very similar among themselves but also quite close to the root of the autoregressive component. This implies the presence of a near common factor in the univariate models for the first differenced series so that the AR and MA roots would almost cancel out. To our knowledge, this feature of the parameterization **MC(b)** used by (and many others Gonzalo [18]) has not been noticed so far in simulation studies on unit roots or cointegration tests. As a consequence of this near common root, the lagged unconditional correlation of Δy_t and Δx_t will tend to be small. For instance, for $\rho = 0.9$, $\sigma = 2$ and $\eta = -0.5$, the first-order unconditional autocorrelations of Δy_t and Δx_t are equal to 0.015 and −0.021, respectively, and the first-order unconditional cross-correlation is equal to −0.032. Similar results are obtained when $\rho = 0.9$ and for different values for σ and η. The first-order unconditional autocorrelations increase very slowly as ρ decreases, for instance, when $\rho = 0.5$, we have 0.05, −0.11 and −0.16 for the first-order unconditional autocorrelation of Δy_t, of Δx_t and the first-order unconditional cross-correlation, respectively.

The empirical size, at a 5% nominal level, for the unit root tests, is reported in Tables 3–6 for the estimated regression without a trend. [4] For each DGP, we test for a unit root both in y_t and in x_t. For each set of parameters in **MC(b)**, in each table, we report the effective size for a fixed value of the "signal-to-noise" ratio and different values of the remaining parameters; for the experiments in **MC(a)**, we report the effective size for the four different parameterizations.

The first general and striking result, common to both parameterization **MC(b)** and **MC(a)**, concerns the presence of important differences in the effective size according to whether y_t or x_t are tested for the presence of a unit root. Considering **MC(b)**, the empirical size increases with η when testing for a unit root in y_t and, on the other hand, it decreases when testing is carried out on x_t. This finding is remarkable and unexpected since the univariate ARMA representations of y_t and x_t share the same AR component and have very similar MA components for most parameterizations. The differences in size can be fully appreciated when $\sigma = 1$ or $\sigma = 2$. For instance, in the case $\sigma = 1, \rho = 0.9, \eta = -0.5$, for most tests based on GLS detrending (but for the MZ_t^{PQ}), the effective size is close to the nominal one when the unit root test is applied to y_t, but it doubles or almost triples

[4] For brevity, our discussion will refer to the model without trend, similar remarks apply when a trend component is included in the regression, simulation results in this case are available upon request from the authors.

when unit root tests are applied to x_t. In addition, the same applies to bootstrap unit root tests. The distortion in the effective size on y_t and x_t is reversed when $\sigma = 1, \rho = 0.9$ but $\eta = 0.5$ and to a lesser extent for smaller values of the AR root such as $\rho = 0.5$. Notice also that only in the case in which $\eta = 0$, do these differences tend to to be negligible.

For the parameterization in **MC(a)**, in Table 6, we continue to observe substantial differences in the effective size according to whether y_t or x_t are being tested for a unit root. These differences are even more pronounced than those in Tables 3–5, perhaps because of the greater range and heterogeneity of the MA component obtained from the parameter values under **MC(a)**. Furthermore, there are noticeable differences among tests and across DGPs: for instance, looking at DGP1, both the DF-GLS and bootstrap tests have reasonable effective sizes for x_t while the effective size more than doubles for the bootstrap tests applied to y_t, but it does not change substantially for the DF-GLS test. Again, for DGP4, the size of the DF-GLS more than doubles when y_t is tested while the size of the bootstrap tests is more than four times larger for y_t than for x_t.

Considering the parameterization in **MC(b)**, from Tables 3–5, we notice that when $\sigma = 0.25$, the empirical size is, in general, close to the nominal one for most tests and it is so, in particular, for the ADF sieve bootstrap unit root tests. In particular, the empirical size in both versions of the ADF boostrap test seem to be more stable and closer to the nominal size than the empirical size of the P_T, MP_T, and DF-GLS tests. However, for $\sigma = 1$, and to a larger extent when $\sigma = 2$, the empirical size of these tests tend to differ more and more from the nominal 5% level. In fact, the effective size increases with σ ranging in the interval $(0.03, 0.09)$ when $\sigma = 0.25$ to the interval $(0.11, 0.25)$ when $\sigma = 2$. Thus, as the variance of the random walk component in y_t and x_t increases, the size of the unit root tests increases, leading to greater size distortion, and the size distortion itself is quite sizeable for $\sigma = 2$, irrespective of ρ and η. In general, GLS detrending tends to increases the empirical size and this exerts a beneficial effect when σ is small, but, on the contrary, it is detrimental for the size when σ is large.

Table 3. Empirical size of unit root tests (no trend)—**MC(b)**, $\sigma = 0.25$.

	$\rho = 0.9$						$\rho = 0.5$					
	$\eta = -0.5$		$\eta = 0$		$\eta = 0.5$		$\eta = -0.5$		$\eta = 0$		$\eta = 0.5$	
Test	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t
						OLS Detrending						
Z_α	0.029	0.058	0.037	0.030	0.044	0.027	0.018	0.054	0.049	0.048	0.065	0.023
MZ_α	0.013	0.025	0.014	0.014	0.018	0.018	0.009	0.029	0.021	0.027	0.032	0.009
Z_t	0.043	0.070	0.047	0.044	0.055	0.046	0.028	0.066	0.061	0.068	0.099	0.051
MZ_t	0.004	0.014	0.005	0.008	0.004	0.003	0.004	0.006	0.014	0.014	0.012	0.002
MSB	0.020	0.037	0.027	0.024	0.035	0.023	0.013	0.039	0.035	0.041	0.041	0.013
ADF	0.038	0.067	0.043	0.040	0.047	0.042	0.029	0.060	0.050	0.057	0.063	0.044
						GLS detrending						
Z_α	0.047	0.084	0.063	0.048	0.088	0.046	0.040	0.077	0.070	0.070	0.098	0.043
MZ_α	0.041	0.071	0.052	0.041	0.077	0.039	0.034	0.07	0.059	0.062	0.083	0.036
MZ_α^{PQ}	0.005	0.009	0.008	0.004	0.007	0.008	0.006	0.012	0.010	0.013	0.008	0.004
Z_t	0.051	0.083	0.062	0.051	0.090	0.052	0.038	0.084	0.073	0.071	0.102	0.044
MZ_t	0.039	0.069	0.045	0.036	0.071	0.039	0.029	0.062	0.050	0.058	0.082	0.034
MSB	0.035	0.061	0.052	0.043	0.066	0.038	0.035	0.068	0.053	0.058	0.069	0.028
P_T	0.030	0.059	0.037	0.033	0.058	0.035	0.027	0.053	0.040	0.052	0.070	0.024
MP_T	0.037	0.067	0.042	0.035	0.071	0.037	0.029	0.061	0.049	0.059	0.082	0.032
$DF-GLS$	0.048	0.078	0.058	0.047	0.081	0.046	0.038	0.069	0.065	0.069	0.090	0.040
						Bootstrap Tests						
ADF_b	0.045	0.058	0.060	0.043	0.064	0.036	0.040	0.051	0.048	0.053	0.062	0.043
ADF_b^r	0.044	0.062	0.062	0.056	0.070	0.040	0.042	0.061	0.059	0.057	0.064	0.044
						Johansen's Trace Test						
J_{tr}	0.722		0.834		0.717		0.059		0.076		0.069	

Table 4. Empirical size of unit root tests (no trend)—**MC(b)**, $\sigma = 1$.

	$\rho = 0.9$						$\rho = 0.5$					
	$\eta = -0.5$		$\eta = 0$		$\eta = 0.5$		$\eta = -0.5$		$\eta = 0$		$\eta = 0.5$	
Test	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t
OLS detrending												
Z_α	0.033	0.098	0.082	0.080	0.107	0.037	0.049	0.146	0.108	0.119	0.164	0.081
MZ_α	0.014	0.057	0.045	0.040	0.062	0.020	0.026	0.090	0.060	0.071	0.087	0.038
Z_t	0.037	0.102	0.082	0.092	0.115	0.049	0.064	0.196	0.149	0.150	0.205	0.087
MZ_t	0.003	0.030	0.017	0.023	0.029	0.009	0.010	0.048	0.029	0.038	0.056	0.016
MSB	0.021	0.077	0.062	0.069	0.089	0.029	0.038	0.109	0.073	0.085	0.112	0.056
ADF	0.034	0.083	0.072	0.082	0.094	0.049	0.056	0.101	0.085	0.088	0.117	0.068
GLS detrending												
Z_α	0.064	0.156	0.137	0.122	0.184	0.069	0.065	0.155	0.126	0.156	0.184	0.100
MZ_α	0.053	0.134	0.118	0.102	0.154	0.056	0.055	0.115	0.100	0.122	0.145	0.078
MZ_α^{PQ}	0.004	0.030	0.019	0.018	0.028	0.007	0.013	0.038	0.031	0.033	0.043	0.016
Z_t	0.061	0.166	0.146	0.127	0.184	0.078	0.075	0.175	0.144	0.168	0.226	0.106
MZ_t	0.046	0.123	0.110	0.096	0.142	0.050	0.055	0.108	0.094	0.116	0.133	0.080
MSB	0.053	0.115	0.103	0.100	0.141	0.048	0.059	0.107	0.089	0.118	0.131	0.070
P_T	0.042	0.107	0.100	0.085	0.121	0.042	0.047	0.094	0.082	0.104	0.109	0.069
MP_T	0.047	0.118	0.111	0.095	0.143	0.049	0.055	0.107	0.093	0.115	0.132	0.076
$DF-GLS$	0.049	0.149	0.126	0.115	0.167	0.066	0.066	0.119	0.113	0.129	0.148	0.091
Bootstrap Tests												
ADF_b	0.049	0.130	0.105	0.091	0.125	0.052	0.067	0.096	0.077	0.079	0.099	0.061
ADF_b^r	0.052	0.141	0.110	0.107	0.136	0.055	0.071	0.099	0.083	0.090	0.109	0.072
Johansen's Trace Test												
J_{tr}	0.710		0.829		0.742		0.056		0.062		0.084	

Table 5. Empirical size of unit root tests (no trend)—**MC(b)**, $\sigma = 2$.

	$\rho = 0.9$						$\rho = 0.5$					
	$\eta = -0.5$		$\eta = 0$		$\eta = 0.5$		$\eta = -0.5$		$\eta = 0$		$\eta = 0.5$	
Test	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t
OLS Detrending												
Z_α	0.097	0.177	0.136	0.134	0.164	0.110	0.195	0.266	0.257	0.245	0.303	0.172
MZ_α	0.061	0.108	0.086	0.086	0.098	0.072	0.095	0.140	0.129	0.134	0.153	0.088
Z_t	0.083	0.159	0.121	0.116	0.153	0.103	0.254	0.423	0.364	0.347	0.432	0.247
MZ_t	0.028	0.053	0.040	0.034	0.052	0.037	0.052	0.101	0.096	0.090	0.106	0.050
MSB	0.079	0.139	0.116	0.115	0.134	0.100	0.120	0.162	0.152	0.160	0.176	0.11 0
ADF	0.072	0.141	0.108	0.100	0.124	0.089	0.125	0.158	0.148	0.149	0.161	0.112
GLS detrending												
Z_α	0.197	0.292	0.233	0.235	0.241	0.195	0.191	0.262	0.219	0.207	0.279	0.189
MZ_α	0.174	0.252	0.211	0.209	0.210	0.164	0.138	0.145	0.149	0.117	0.163	0.116
MZ_α^{PQ}	0.035	0.047	0.037	0.030	0.035	0.037	0.044	0.076	0.069	0.060	0.086	0.04
Z_t	0.209	0.297	0.246	0.235	0.256	0.214	0.240	0.381	0.307	0.279	0.369	0.238
MZ_t	0.157	0.234	0.192	0.197	0.194	0.164	0.134	0.130	0.138	0.116	0.155	0.106
MSB	0.157	0.237	0.187	0.200	0.205	0.150	0.118	0.130	0.138	0.109	0.150	0.109
P_T	0.137	0.200	0.171	0.172	0.170	0.138	0.119	0.112	0.122	0.099	0.132	0.089
MP_T	0.154	0.229	0.188	0.197	0.192	0.158	0.133	0.132	0.138	0.112	0.154	0.103
$DF-GLS$	0.188	0.271	0.223	0.221	0.235	0.199	0.149	0.168	0.156	0.126	0.176	0.122
Bootstrap Tests												
ADF_b	0.194	0.223	0.218	0.193	0.238	0.169	0.118	0.136	0.139	0.124	0.148	0.113
ADF_b^r	0.200	0.239	0.226	0.206	0.251	0.188	0.129	0.149	0.135	0.130	0.153	0.121
Johansen's Trace Test												
J_{tr}	0.730		0.849		0.711		0.070		0.052		0.059	

Table 6. Empirical size of unit root tests (no trend)—**MC(a)**.

	DGP1		DGP2		DGP3		DGP4	
Test	y_t	x_t	y_t	x_t	y_t	x_t	y_t	x_t
				OLS Detrending				
Z_α	0.927	0.086	0.283	0.016	0.140	0.739	0.672	0.041
MZ_α	0.084	0.024	0.112	0.009	0.017	0.020	0.188	0.011
Z_t	0.994	0.161	0.481	0.032	0.272	0.933	0.899	0.109
MZ_t	0.070	0.010	0.075	0.003	0.010	0.012	0.154	0.003
MSB	0.086	0.033	0.123	0.013	0.022	0.022	0.208	0.023
ADF	0.136	0.056	0.134	0.03	0.048	0.072	0.228	0.037
				GLS detrending				
Z_α	0.338	0.078	0.270	0.036	0.062	0.273	0.216	0.033
MZ_α	0.012	0.050	0.106	0.034	0.029	0.006	0.058	0.031
MZ_α^{PQ}	0.033	0.006	0.053	0.004	0.003	0.006	0.065	0.003
Z_t	0.559	0.094	0.393	0.034	0.096	0.490	0.343	0.029
MZ_t	0.012	0.047	0.101	0.032	0.028	0.006	0.057	0.028
MSB	0.010	0.046	0.093	0.031	0.032	0.005	0.052	0.031
P_T	0.012	0.035	0.085	0.029	0.027	0.006	0.054	0.024
MP_T	0.013	0.046	0.100	0.032	0.029	0.006	0.057	0.028
$DF-GLS$	0.062	0.059	0.133	0.028	0.051	0.052	0.093	0.039
				Bootstrap Tests				
ADF_b	0.136	0.052	0.104	0.039	0.061	0.085	0.189	0.044
ADF_b^r	0.122	0.051	0.113	0.034	0.056	0.078	0.188	0.041
				Johansen's Trace Test				
J_{tr}	0.044		0.064		0.041		0.037	

Considering the four parameterizations under **MC(a)**, the behavior of unit root tests is more heterogeneous, as is the pattern of AR and MA roots. Under OLS detrending, the Z_α and Z_t tests have large size distortion for all parameterizations. The modifications suggested by Perron and Ng [2] are somehow effective in reducing the distortion, but the behavior of the M tests is not stable across DGPs, and the same remark applies to the ADF test. Under GLS detrending, the DF-GLS test by Elliott *et al.* [13] has the best performance, showing an effective size very close to the nominal one in all cases but for y_t in DGP2 and DGP4 where, in fact, the MA component has a coefficient close to -1. The bootstrap unit root tests have greater size distortion than the DF-GLS but, for y_t in DGP2 and DGP4, exactly in those cases where the DF-GLS test does not perform well.

Finally, we consider consider Johansen's trace test under the null of one cointegrating vector. For parameters in **MC(b)**, the test statistics are severely biased when $\rho = 0.9$, irrespective of the values taken by σ and η, while it has a size close to the nominal one for $\rho = 0.05$ and its performance, in the latter case, is superior to those of the standard and bootstrap unit root tests. For the parameterizations in **MC(a)**, Johansen's trace test also displays a very good behavior since about 5% of the time, we reject no cointegration in favor of stationarity of the VAR in (1) in all cases. However, we should bear in mind that the AR root is rather small now and that, from the previous results based on **MC(b)**, Johansen's test is adversely affected by large values of ρ.

We do not consider the power properties of the unit root tests considered here. Ng and Perron [14] provide simulation evidence that the DF-GLS has better power then the M tests, even though the latter have better size properties. For the bootstrap unit root tests, we take the results of the extensive simulation study by Palm *et al.* [16] who found that the ADF sieve bootstrap test performs better under a variety of DGPs with and without an MA component. An extensive simulation study on the power properties of the tests considered here for univariate time series generated by a cointegrated VAR is left for a further investigation.

4. Conclusions

A standard practice in cointegration analysis is to run unit root tests separately for each single time series in the multivariate system. However, univariate time series are most often observed as a part of a more general multivariate model. It turns out that the univariate models implied by a VAR data generating process always have a finite order MA component (e.g., see [6]). This feature may explain why an MA component has often been found in univariate ARMA models for economic time series and, given the well-known size distortion of popular unit root tests in the presence of a large negative coefficient in the MA component, it has important implications for unit root tests in univariate settings. In a small simulation experiment under cointegration, we find that (a) there can be substantial differences in the size distortion according to whether the unit root test is applied on y_t or x_t and that this occurs for the ADF sieve bootstrap unit root test too; (b) most tests perform well when the "signal-to-noise" ratio is small but the size distortion can be large when the "signal-to-noise" ratio increases; (c) the ADF sieve bootstrap unit root tests are not immune from size distortion; and (d) Johansen's trace test based on the VAR model exhibits great size distortion when the root of the AR component is large.

Acknowledgments: We would like to thank three anonymous referees and the editor for very helpful comments.

Author Contributions: Both authors contributed equally to the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schwert, W. Testing for unit roots: a montecarlo investigation. *J. Bus. Econ. Stat.* **1989**, *7*, 147–159.
2. Perron, P.; Ng, S. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Rev. Econ. Stud.* **1996**, *63*, 435–563.
3. Reed, W.R. *Unit Root Tests, Size Distortion, and Cointegrated Data*; Working Paper; University of Canterbury: Christchurch, New Zealand, 2014.
4. Reed, W.R. *Univariate Unit root Tests Perform Poorly When Data are Cointegrated*; Working Paper; University of Canterbury: Christchurch, New Zealand, 2016.
5. Zellner, A.; Palm, F. Time series analysis and simultaneous equation econometric models. *J. Econom.* **1974**, *2*, 17–54.
6. Tsay, R.S. *Analysis of Financial Time Series*, 3rd ed.; John Wiley & Sons: New York, NY, USA, 2010.
7. Cubadda, G.; Hecq, A.; Palm, F. Studying co-movements in large multivariate data prior to multivariate modelling. *J. Econom.* **2009**, *148*, 25–35.
8. Cubadda, G.; Triacca, U. An alternative solution to the autoregressive paradox in time series analysis. *Econ. Model.* **2011**, *28*, 1451–1454.
9. Maravall, A.; Mathis, A. Encompassing univariate models in multivariate time series. *J. Econom.* **1994**, *61*, 197–233.
10. Said, S.E.; Dickey, D.A. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* **1984**, *71*, 599–607.
11. Phillips, P.C.B.; Perron, P. Testing for a unit root in time series regression. *Biometrika* **1987**, *75*, 335–346.
12. Stock, J.H. A Class of Tests for Integration and Cointegration. In *Cointegration, Causality and Forecasting: A Festschrift for Clive W.J. Granger*; Oxford University Press: Oxford, UK, 1999; pp. 135–167.
13. Elliott, G.; Rothenberg, T.J.; Stock, J.H. Efficient tests for an autoregressive unit root. *Econometrica* **1996**, *64*, 813–836.
14. Ng, S.; Perron, P. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* **2001**, *69*, 1519–1554.
15. Perron, P.; Qu, Z. A simple modification to improve the finite sample properties of ng and perron's unit root tests. *Econ. Lett.* **2007**, *94*, 12–19.
16. Palm, F.C.; Smeekes, S.; Urbain, J. Bootstrap unit root tests: Comparisons and extensions. *J. Time Ser. Anal.* **2008**, *29*, 371–401.
17. Chang, Y.; Park, J.Y. A sieve bootstrap for the test of a unit root. *J. Time Ser. Anal.* **2003**, *24*, 379–400.

18. Gonzalo, J. Five alternative methods of estimating long-run equilibrium relationships. *J. Econom.* **1994**, *60*, 203–233.
19. Psaradakis, Z. Bootstrap tests for an autoregresive unit root in the presence of weakly dependent errors. *J. Time Ser. Anal.* **2001**, *22*, 577–594.

© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).