Multivariate Unit Root Tests

R. G. Flôres Jr., P-Y. Preumont and A. Szafarz

This paper presents a new multivariate test for the detection of unit roots. Use is made of the possible correlations between the disturbances of different series, and constrained and unconstrained SURE estimators are employed. The corresponding asymptotic distributions are obtained and a table with a few critical values, for the case of two series, is generated. Some simulations indicate that the procedure performs better than the existing alternatives.

JEL Classifications: C12, C30, C52, C33, C32
Keywords: unit root, multivariate test, stationarity, panel data

CEB Working Paper N° 95/001
January 1995
MULTIVARIATE UNIT ROOT TESTS

January 1995

Renato G. Flôres Jr., Pierre-Yves Preumont and Ariane Szafarz

Université Libre de Bruxelles
C.P. 145/1
50, avenue F. D. Roosevelt
1050 Bruxelles
Belgium

Abstract

This paper presents a new multivariate test for the detection of unit roots. Use is made of the possible correlations between the disturbances of different series, and constrained and unconstrained SURE estimators are employed. The corresponding asymptotic distributions are obtained and a table with a few critical values, for the case of two series, is generated. Some simulations indicate that the procedure performs better than the existing alternatives.

Keywords: unit root, multivariate test, stationarity, panel data
JEL codes: C12, C30, C52, C33, C32

1We thank Pierre Perron, Philippe Jorion and Khalid Sekkat for comments, suggestions and references.
I. INTRODUCTION.

Since Dickey and Fuller’s (1979, 1981) benchmark contribution, many papers have stressed the statistical problems related to unit root tests. Phillips and Perron (1988) and Ng and Perron (1995) are only two examples of attempts to shed some light on unwanted behaviour - under different settings - and of corrections or new procedures to deal with them. Given the considerable diversity of the alternative hypotheses this is not an easy task, and sometimes general statements - not very encouraging for the daily practitioner - are produced. As an example, there is fairly general agreement that moving averages are the most dangerous alternatives. However, though Hall (1989)'s suggestion theoretically solves the problem, the issue of identifying in practice the existence of the moving average remains open. Moreover, as unit root tests are a prerequisite for other procedures, like cointegration testing, this lack of power may create an embarrassing pre-testing bias (see, for instance, Elliot (1993)).

In all the studies up to now attention has been centred on the case of one series. However, the need to perform the test on a group of series closely related, like exchange-rates for several currencies, or basic macroeconomic indexes, calls for a multivariate approach to the problem. Indeed, one could take advantage of the correlations between the residuals in the different series and, like in Zellner (1962)'s pioneering article, improve the efficiency of the estimators and tests. Abuaf and Jorion (1990) have followed this direction, in the context of real exchange rate series. They improved the estimation of the autocorrelation coefficients though constraining a common dynamics under the null. Quah (1992) and Levin and Lin (1993) have approached the problem in a somewhat different way, viewing the multivariate framework as a panel structure in which at least one of the dimensions goes to infinity.

In this paper we develop multivariate testing procedures generalising Abuaf and Jorion (1990)'s idea. The cases of a "common" unit root and of different nonstationary/stationary behaviour are addressed and the asymptotic distributions are found. Critical values for the case of two series are presented, and some preliminary simulations point favourably to the proposed tests.

The general model and related tests are discussed in the next section. Critical values and simulations are presented in Section 3, while Section 4 concludes. The derivation of the limit distributions is presented in an Appendix.
II. THE MODEL AND THE TESTS.

We consider the following set of series:

\[ x_{it} = \alpha_i x_{it-1} + \nu_{it}, \quad x_{t0} \equiv 0, \quad 1 \leq i \leq p, \]

where the vector process \( \{\nu_i\} = \{\nu_{it}\} \) satisfies the hypotheses in Phillips and Solo (1992).

Several hypotheses of the form:

\[ H_0 : \alpha_i = 1 \text{ for } i \in I \ (I \subset \{1, ..., p\}), \quad \alpha_i < 1 \text{ for } i \in I^C \]  

(2.1)

can be of interest. In particular, the case:

\[ H_0 : \alpha_i = 1, \quad 1 \leq i \leq p, \]

(2.2)

imposes a common I (1) dynamics to all series.

We shall consider tests of the above unit root hypotheses in which the autoregressive coefficients are estimated by means of Zellner's SURE method. The equalities present either in (2.1) or (2.2) can be taken into account or not in the estimation procedure, giving rise to what we shall call constrained and unconstrained tests. The former, in the case (2.2), have been proposed and numerically studied by Abuaf and Jorion (1990) and will be called the AJ test.

If the parameter \( \alpha = (\alpha_i) \) is estimated by the method SURE, then (2.2) represents a limit situation in which the equalities in the null may be fully used. However, rejection of the null may have many forms; actually, any of those represented in (2.1).

Our procedure starts then with the stronger condition (2.2). As will be seen in the next section, this does not necessarily imply that the constrained AJ test should be used. If the null is rejected, inspection of the estimated \( \hat{\alpha}_i \) may help in choosing an appropriate version of (2.1) as the new null. In this case, the fact that some \( \{x_{it}\} \) series are stationary is crucial and may greatly improve the results. Adequate testing of (possible) successive steps may eventually lead to a final structure of stationary and
non-stationary series. This can produce a feasible methodology even for moderately high values of \( p \).

The asymptotic distributions for the three types of tests, when \( p = 2 \), are now presented. All proofs are given in the Appendix.

Consider the bivariate model:

\[
\begin{align*}
    x_t &= \alpha_1 x_{t-1} + u_t, \\
    y_t &= \alpha_2 y_{t-1} + v_t, \\
    x_0 = y_0 &= 0,
\end{align*}
\]

where \( \{ [u_t \ v_t] \}_{t=0}^{\infty} \) is an i.i.d. zero mean process\(^2\), with a regular variance-covariance matrix \( \Omega \), given by:

\[
\Omega = \begin{bmatrix}
    \sigma_{11} & \sigma_{12} \\
    \sigma_{12} & \sigma_{22}
\end{bmatrix},
\]

and the corresponding correlation is

\[
\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}. \tag{2.5}
\]

Let us suppose first that \( \Omega \) is known. The following tests are the bivariate versions of (2.2) and (2.1), respectively:

\[
H_0^{(1)} : \alpha_1 = 1, \ \alpha_2 = 1 \tag{2.6}
\]

and

\[
H_0^{(2)} : \alpha_1 = 1, \ \alpha_2 < 1. \tag{2.7}
\]

Hypothesis \( H_0^{(1)} \) can be dealt with in two different ways. We shall first discuss the distribution of the test statistic numerically studied by Abuaf and Jorion (1990), in which the equality constraint is used for estimating the autoregressive coefficient. Then, hypothesis \( H_0^{(1)} \) will be studied under an unconstrained SURE estimator.

As proved in the Appendix, the limit distribution of the AJ test is:

\[^2\text{More general hypotheses on the disturbances can be formulated, following Phillips and Solo (1992).}\]
\[ T(\hat{\alpha} - 1) \Rightarrow \frac{\int_0^1 W_u(s)dW_u(s) - \rho_{12} \left[ W_u(1)W_v(1) - \rho_{12} \right] + \int_0^1 W_v(s)dW_v(s)}{\frac{1}{2} \left( W_u^2(1) + W_v^2(1) \right) - 1}. \]  
(2.8)

where \( \{W_u(s)\}_{s \in [0,1]} \) and \( \{W_v(s)\}_{s \in [0,1]} \) are standard brownian motions associated respectively to \( \left\{ \frac{u_i}{\sqrt{\sigma_{11}}} \right\} \) and \( \left\{ \frac{v_i}{\sqrt{\sigma_{22}}} \right\} \).

Formula (2.8) shows that only the correlation coefficient appears in the limit distribution. This means that if the unknown variances are estimated in a consistent way, this limit will be unchanged.

Two extreme cases should be mentioned. On the one hand, if \( \rho_{12} = 0 \), i.e. no advantage exists in the SURE procedure, an expression close to the DF classical test is found:

\[ T(\hat{\alpha} - 1) \Rightarrow \frac{\int_0^1 W_u(s)dW_u(s) + \int_0^1 W_v(s)dW_v(s)}{\int_0^1 \left( W_u^2(s) + W_v^2(s) \right) ds} = \frac{1}{2} \left( W_u^2(1) + W_v^2(1) \right) - 1. \]  
(2.9)

On the other hand, when \( \rho_{12} \to 1 \), expression (2.8) tends to:

\[ T(\hat{\alpha} - 1) \Rightarrow \frac{\int_0^1 \left( W_u(s) - W_v(s) \right) dW_u(s) + \int_0^1 \left( W_v(s) - W_u(s) \right) dW_v(s)}{\int_0^1 (W_u(s) - W_v(s))^2 ds} = \frac{1}{2} \left( W_u(1) - W_v(1) \right)^2. \]  
(2.10)

what should really be taken as a limit case, since, if \( \rho_{12} = 1 \), the variance-covariance matrix \( \Omega \) is not invertible.

If unconstrained SURE estimation is used under hypothesis \( H_0^{(1)} \), the limit distribution of the estimator of the autoregressive coefficient \( \alpha_i \) in (2.3), for instance, will be:
\[
T(\hat{\alpha}_1 - 1) \Rightarrow \frac{\int_0^1 W_u(s) \left\{ W_u(s) \left[ \int_0^1 W_u dW_u - \rho_{12} \int_0^1 W_u dW_v \right] + \rho_{12} W_u(s) \left[ \int_0^1 W_u dW_v - \rho_{12} \int_0^1 W_u dW_u \right] \right\} ds}{\left( \int_0^1 W_u^2(s) ds \right) \left( \int_0^1 W_v^2(s) ds \right) - \rho_{12} \left( \int_0^1 W_u(s) W_v(s) ds \right)^2}
\]  

(2.11)

If \( \rho_{12} = 0 \), we have exactly the Dickey-Fuller statistic:

\[
T(\hat{\alpha} - 1) \Rightarrow \frac{\int_0^1 W_u dW_u}{\int_0^1 W_u^2(s) ds}.
\]

We now consider hypothesis \( H_0^{(2)} \). The asymptotic distribution for \( T(\hat{\alpha}_1 - 1) \) will be:

\[
T(\hat{\alpha}_1 - 1) \xrightarrow{f} \frac{\int_0^1 W_u(s) dW_u(s) - \rho_{12} \int_0^1 W_u(s) dW_v(s)}{\int_0^1 W_u^2(s) ds}.
\]  

(2.12)

The Dickey-Fuller statistic is again found if \( \rho_{12} = 0 \), while the case when \( \rho_{12} \to 1 \) would give the maximal gain achieved from the correlation between the disturbances. This limit expression is:

\[
T(\hat{\alpha}_1 - 1) \xrightarrow{f} \frac{\int_0^1 W_u(s) dW_u(s) - \int_0^1 W_u(s) dW_v(s)}{\int_0^1 W_u^2(s) ds}.
\]  

(2.13)

Notice that the distribution in (2.13) is independent of the true \( \alpha_2 \) value, only its stationarity being relevant. This means that, for sufficiently large samples, tables of critical values do not need to take into account this parameter. Actually, the stationary coefficient "helps" in weakening the influence of the (other) non-stationary series. To illustrate this statement, consider the model:

\[
\begin{align*}
    x_t &= x_{t-1} + u_t \\
y_t &= \alpha_2 y_{t-1} + v_t, \quad |\alpha_2| < 1
\end{align*}
\]

where \( E[u_t, v_t] = \delta_{u, k} \)  

(2.14)
One can successively write:

\[ \text{cov}(x_t, y_t) = \alpha_2 \text{cov}(x_{t-1}, y_{t-1}) + k = \alpha'_2 \text{cov}(x_0, y_0) + k \sum_{i=0}^{t-1} \alpha'_2, \]

so that as \( t \to \infty \) the covariance approaches the finite value \( k \frac{1}{1 - \alpha_2} \).

The results in this section emphasize that the correlations between the disturbances of the series of interest play a crucial role. Unfortunately, these correlations are unknown in the majority of the practical cases. Nevertheless, as with the feasible SURE, use of a consistent estimator for them guarantees the good properties of the theoretical test statistics. This issue is further supported in the next section.

III. MONTE CARLO RESULTS


In order to have an idea of the behaviour of the proposed tests we have first compared the univariate use of the DF test with the following possibilities:

i) the AJ procedure with known and unknown correlation (to be called AJK and AJU, respectively);

ii) the non-constrained approach to (2.2) with known and unknown correlation (to be called MK and MU, respectively).

Table 1 gives, for \( p=2 \), the critical 1% and 5% values for the statistic \( T(\hat{\alpha}_1 - 1) \), in the cases \( T=100 \) and \( T=500 \). The figures were generated by 5000 replications, with the residual process having unit variances and correlation coefficients of \( \pm 0.5 \) and \( \pm 0.9 \). A simple symmetry argument - supported by the formulae in Section 2 - makes it necessary to show only the results for positive correlations. To give an idea of the accuracy of the Monte Carlo simulations, DF values were also generated and are displayed in the first row of each sample size.
Table 1. Critical 1% and 5% values for $T(\hat{\alpha}_i - 1)$.

<table>
<thead>
<tr>
<th>Sample size : $T = 100$</th>
<th>1%</th>
<th>5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>correlation coefficient</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>DF</td>
<td>-13.3</td>
<td>-7.9</td>
</tr>
<tr>
<td>AJK</td>
<td>-6.9</td>
<td>-6.8</td>
</tr>
<tr>
<td>MK</td>
<td>-11.7</td>
<td>-8.1</td>
</tr>
<tr>
<td>AJU</td>
<td>-7.0</td>
<td>-6.9</td>
</tr>
<tr>
<td>MU</td>
<td>-11.9</td>
<td>-8.2</td>
</tr>
<tr>
<td>M* (0.5)</td>
<td>-11.6</td>
<td>-5.5</td>
</tr>
<tr>
<td>(0.9)</td>
<td>-11.7</td>
<td>-7.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample size : $T = 500$</th>
<th>1%</th>
<th>5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>correlation coefficient</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>DF</td>
<td>-13.7</td>
<td>-8.0</td>
</tr>
<tr>
<td>AJK</td>
<td>-7.1</td>
<td>-7.1</td>
</tr>
<tr>
<td>MK</td>
<td>-11.8</td>
<td>-8.5</td>
</tr>
<tr>
<td>AJU</td>
<td>-7.1</td>
<td>-7.1</td>
</tr>
<tr>
<td>MU</td>
<td>-11.9</td>
<td>-8.6</td>
</tr>
<tr>
<td>M* (0.5)</td>
<td>-11.1</td>
<td>-4.2</td>
</tr>
<tr>
<td>(0.9)</td>
<td>-11.7</td>
<td>-5.2</td>
</tr>
</tbody>
</table>

For all the four tests, the increase from 100 to 500 observations does not change much the critical values. Also, as suggested by the asymptotic results, there is practically no variation whether the disturbances correlation is known or consistently estimated. On the other hand, though the magnitude of the correlation itself does not affect much the AJ abscissas, it has an important impact on the non-constrained M version. This, combined with the fact that AJ's values are always stricter, simply reflects what has been mentioned in the previous section. By using the very hypothesis in the estimation procedure, the constrained AJ tests surely produce stricter critical values under the null. In the unconstrained M versions, the higher the residuals correlation the more one gains with the multivariate procedure and the closer one moves to the constrained values.

The table is completed with two additional rows, denoted by M*, for each sample size. These rows give the critical values corresponding to the case where the
second root is stationary, under known variances and covariances. Though the asymptotic distribution is here independent of the particular value of the stationary coefficient, two values (0.5 and 0.9) were tried to check the behaviour in finite samples. Again, the magnitude of the correlation plays a central role. This effect is reasonably less pronounced at T=500.

It is also worth noticing that, for all tests except M*, the critical values increase slightly with the sample size because of the non-stationarity of the simulated processes. The opposite is true with M* since the stationary component attenuates the asymptotic impact (see the end of Section 2).

3.2. Power.

Violations of the null can take place in many ways. To check the performance of the different tests, the null (2.2) was violated by alternatives with one stationary coefficient of respectively 0.50, 0.90 and 0.97, supposing, again, residuals correlations of 0.5 and 0.9. For each case, 1000 runs were made.

Table 2 shows the percentage of correct diagnoses for the constrained (AJU) and unconstrained (MU) tests, with the values obtained with the univariate DF - test shown to ease the comparison.

Now, the correlation matters significantly for the constrained test as it acts in the same direction as the $\alpha_2$ values. Indeed, with a low $\alpha_2$, the resulting estimate is a compromise between both values and the percentage of incorrect diagnoses is high. However, it is less bad if the correlation is higher, as fuller use of the estimation technique is made.
Table 2: Percentages of correct diagnoses for $\alpha_1$.

<table>
<thead>
<tr>
<th>$\alpha_1 = 1, \alpha_2 = 0.50$</th>
<th>DF</th>
<th>AJU</th>
<th>MU</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>94.7</td>
<td>69.4</td>
<td>95.5</td>
</tr>
<tr>
<td>0.90</td>
<td>95.2</td>
<td>85.7</td>
<td>99.0</td>
</tr>
<tr>
<td>$\alpha_1 = 1, \alpha_2 = 0.90$</td>
<td>DF</td>
<td>AJU</td>
<td>MU</td>
</tr>
<tr>
<td>0.50</td>
<td>95.3</td>
<td>77.3</td>
<td>95.3</td>
</tr>
<tr>
<td>0.90</td>
<td>94.6</td>
<td>93.1</td>
<td>97.3</td>
</tr>
<tr>
<td>$\alpha_1 = 1, \alpha_2 = 0.97$</td>
<td>DF</td>
<td>AJU</td>
<td>MU</td>
</tr>
<tr>
<td>0.50</td>
<td>94.7</td>
<td>86.4</td>
<td>95.1</td>
</tr>
<tr>
<td>0.90</td>
<td>94.0</td>
<td>97.3</td>
<td>94.8</td>
</tr>
</tbody>
</table>

On the other hand, the unconstrained MU test performs rather well, with a slight tendency to overaccept. In any case, it is much more robust than the AJU's.

The situation gets worse when we turn to the stationary coefficient estimation. As Table 3 shows, the AJU test, as expected, performs very badly. The MU version performs very well until 0.90, and still remains superior to the other two at the already misleading value of 0.97.

Table 3: Percentage of correct diagnoses for $\alpha_2$.

<table>
<thead>
<tr>
<th>$\alpha_1 = 1, \alpha_2 = 0.50$</th>
<th>DF</th>
<th>AJU</th>
<th>MU</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>100.0</td>
<td>30.7</td>
<td>100.0</td>
</tr>
<tr>
<td>0.90</td>
<td>100.0</td>
<td>14.3</td>
<td>100.0</td>
</tr>
<tr>
<td>$\alpha_1 = 1, \alpha_2 = 0.90$</td>
<td>DF</td>
<td>AJU</td>
<td>MU</td>
</tr>
<tr>
<td>0.50</td>
<td>77.7</td>
<td>22.7</td>
<td>88.6</td>
</tr>
<tr>
<td>0.90</td>
<td>76.6</td>
<td>6.9</td>
<td>99.6</td>
</tr>
<tr>
<td>$\alpha_1 = 1, \alpha_2 = 0.97$</td>
<td>DF</td>
<td>AJU</td>
<td>MU</td>
</tr>
<tr>
<td>0.50</td>
<td>15.6</td>
<td>13.6</td>
<td>22.4</td>
</tr>
<tr>
<td>0.90</td>
<td>16.2</td>
<td>2.7</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Summing up, there is an evident advantage in performing the multivariate tests. Moreover, contrary to the basic intuition, the unconstrained version should be used in the first step of the procedure.

---

3Actually, the AJU column equals, but for a round-off error, 100 minus the corresponding column value in Table 2.
IV. CONCLUSIONS.

Exploration of the multivariate structure of the disturbances of related series produces fruitful results in unit root testing. Actually, the unconstrained estimation proposed in this paper seems to represent an adequate compromise between the classical Dickey-Fuller approach and the multivariate constrained methodology suggested by Abuaf and Jorion. From the former, it keeps the flexibility of individually testing the coefficients while, as in the latter, it exploits the correlations between the series.

This paper has provided some preliminary evidence of the possible advantages of the proposition and produced some estimates of a few critical values. Further work is needed in order to generate a complete set of tables for practical use. Moreover, a deeper power study, investigating the behaviour under problematic error structures like moving-averages is needed to fully certify the proposal.
Appendix: Proofs of the asymptotic results.

Proof of (2.8)

If \( \alpha_1 = \alpha_2 = \alpha \), model (2.3) can be rewritten as:

\[
\begin{align*}
\begin{cases}
    x_i = \alpha x_{i-1} + u_i, \\
y_i = \alpha y_{i-1} + v_i,
\end{cases}
x_0 = y_0 = 0.
\end{align*}
\]

The inverse of the matrix \( \Omega \) will be denoted:

\[
\Omega^{-1} = \begin{bmatrix}
\sigma_{11} & \sigma_{12} \\
\sigma_{12} & \sigma_{22}
\end{bmatrix}.
\]

If \( \alpha \) in (A.1) is estimated by a SURE procedure under the equality constraint, given observations \( \{x_t, y_t\}_{t=1}^T \), the formula for the estimator will be:

\[
\hat{\alpha} = \frac{\sigma_{11} \sum x_{t-1} x_t + \sigma_{12} \left( \sum x_{t-1} y_t + \sum y_{t-1} x_t \right) + \sigma_{22} \sum v_{t-1} v_t}{\sigma_{11} \sum x_{t-1}^2 + 2 \sigma_{12} \sum x_{t-1} v_{t-1} + \sigma_{22} \sum v_{t-1}^2}.
\]

Given model (A.1), under the null one gets:

\[
\hat{\alpha} = \frac{\sigma_{11} \sum x_{t-1} u_t + \sigma_{12} \left( \sum x_{t-1} y_t + \sum y_{t-1} u_t \right) + \sigma_{22} \sum v_{t-1} v_t}{\sigma_{11} \sum x_{t-1}^2 + 2 \sigma_{12} \sum x_{t-1} v_{t-1} + \sigma_{22} \sum v_{t-1}^2}
\]

Dividing the numerator and denominator in the RHS of (A.3) by \( (T-1)^2 \), and taking into account the links between the elements of \( \Omega \) and \( \Omega^{-1} \), standard functional central limit results for I(1) processes lead to:

\[
T(\hat{\alpha} - 1) \Rightarrow \left( \lim \limits_{T \to \infty} \frac{T}{T-1} \right) \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \frac{1}{\sigma_{11}} \int_0^1 W_u(s) dW_u(s) - \sigma_{12} \sqrt{\sigma_{11} \sigma_{22}} \left[ W_u(l) W_v(l) - \rho_{12} \right] + \sigma_{11} \sigma_{22} \frac{1}{\sigma_{11}^2} \int_0^1 W'_u(s) dW'_u(s) \\
\sigma_{12} & \sigma_{22} & \frac{1}{\sigma_{12}} \int_0^1 W_u(s) dW_u(s) - \sigma_{12} \sqrt{\sigma_{11} \sigma_{22}} \left[ W_u(l) W_v(l) - \rho_{12} \right] + \sigma_{11} \sigma_{22} \frac{1}{\sigma_{12}^2} \int_0^1 W'_u(s) dW'_u(s)
\end{bmatrix}
\]

\[
\left[ W_u(s) ds - 2 \sigma_{12} \sqrt{\sigma_{11} \sigma_{22}} \int_0^1 W_u(s) W_v(s) ds + \sigma_{11} \sigma_{22} \int_0^1 W'_u(s) dsight]
\]
where \( \{W_u(s)\}_{s \in [0,1]} \) and \( \{W_v(s)\}_{s \in [0,1]} \) are standard brownian motions associated respectively to \( \left\{ \frac{u_t}{\sqrt{\sigma_{11}}} \right\} \) and \( \left\{ \frac{v_t}{\sqrt{\sigma_{22}}} \right\} \).

After simplification, (A.4) leads to the limit distribution in (2.8).

**Proof of (2.11)**

The unconstrained estimator of the autoregressive coefficient \( \alpha_1 \) in (2.3) will be:

\[
\hat{\alpha}_1 = \frac{\sigma_{22} \left( \sum_{t=1}^{T} y_{t-1}^2 \right) \left[ \sigma_{11} \sum_{t=1}^{T} x_{t-1} y_t + \sigma_{12} \sum_{t=1}^{T} x_{t-1} y_{t-1} \right] - \sigma_{12} \left( \sum_{t=1}^{T} x_{t-1} y_t \right) \left[ \sigma_{12} \sum_{t=1}^{T} y_{t-1} y_t + \sigma_{22} \sum_{t=1}^{T} y_{t-1} y_{t-1} \right]}{\sigma_{11} \sigma_{22} \left( \sum_{t=1}^{T} x_{t-1}^2 \right) \left( \sum_{t=1}^{T} y_{t-1}^2 \right) - (\sigma_{12})^2 \left( \sum_{t=1}^{T} x_{t-1} y_{t-1} \right)^2}.
\]

Under the null, this will lead to:

\[
\hat{\alpha}_1 - 1 = \frac{\sigma_{22} \left( \sum_{t=1}^{T} y_{t-1}^2 \right) \left[ \sigma_{11} \sum_{t=1}^{T} x_{t-1} u_t + \sigma_{12} \sum_{t=1}^{T} x_{t-1} v_{t-1} \right] - \sigma_{12} \left( \sum_{t=1}^{T} x_{t-1} y_t \right) \left[ \sigma_{12} \sum_{t=1}^{T} y_{t-1} u_t + \sigma_{22} \sum_{t=1}^{T} y_{t-1} v_{t-1} \right]}{\sigma_{11} \sigma_{22} \left( \sum_{t=1}^{T} x_{t-1}^2 \right) \left( \sum_{t=1}^{T} y_{t-1}^2 \right) - (\sigma_{12})^2 \left( \sum_{t=1}^{T} x_{t-1} y_{t-1} \right)^2}.
\]

After dividing both members of the RHS by \( (T-1)^4 \), use of the asymptotic theorems and of the relations between the elements of \( \Omega \) et \( \Omega^{-1} \) gives the result.

**Proof of (2.12)**

The stationary hypothesis implies that:

\[
(A.7) \ \text{plim} \frac{1}{T} \sum_{t=1}^{T} y_{t-1} v_t = \text{plim} \frac{1}{T} \sum_{t=1}^{T} y_{t-1} u_t = 0, \quad \text{plim} \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 = \sigma_2 \frac{1}{1-\alpha_2}.\]

Moreover, as \( \frac{1}{T-1} \sum_{t=1}^{T} x_{t-1} y_{t-1} \) weakly converges to a non-degenerate variable, \( \frac{1}{(T-1)^2} \sum_{t=1}^{T} x_{t-1} y_{t-1} \) converges in probability to zero. The result follows then from (A.5), after suitable division by \( (T-1)^4 \).
REFERENCES


Elliot G. (1993), "On the Robustness of Cointegration Methods when Regressors Almost have Unit Roots", Harvard University, mimeo.


