Dissertação de Mestrado

Comparing Two Multivariate Unit Root Tests

Guilherme N. L. Veiga da Rocha
EPGE - Fundação Getulio Vargas

Orientador: Renato G. Flôres Jr. (EPGE / FGV-RJ)

\textsuperscript{1}rocha@fgvmail.br
Abstract

In this essay, a method for comparing the asymptotic power of the multivariate unit root tests proposed in Phillips & Durlauf (1986) and Flöres, Preumont & Szafarz (1996) is proposed. In order to determine the asymptotic power of the tests the asymptotic distributions under the null hypothesis and under the set of alternative hypotheses described in Phillips (1988) are determined.

In addition, a test which combines characteristics of both tests is proposed and its distributions under the null hypothesis and the same set of alternative hypotheses are determined. This allows us to determine what causes any difference in the asymptotic power of the two tests against the set of alternative hypotheses considered.
Agradecimentos

Primeiramente, gostaria de agradecer ao meu orientador, professor Renato Flores. O seu profundo conhecimento sobre diversas áreas aliado à sua grande disposição para ensinar certamente contribuíram para tornar os meus anos na EPGE extremamente proveitosos, além de terem sido um enorme incentivo a prosseguir na carreira acadêmica. Agradeço também ao professor Marcelo Fernandes pela paciência e inabalável disposição para discutir dúvidas e ideias, assim como pelas sugestões dadas na realização desse trabalho. Agradeço também aos professores Paulo Klinger e Aloísio Araújo por terem incrementado tanto os meus conhecimentos como o meu interesse pela Matemática. Aos demais professores da EPGE, agradeço pela oportunidade de conviver por três anos em um ambiente intelectualmente rico.

Esses agradecimentos não estariam completos sem a presença dos amigos Frederico Estrella Carneiro Valladares, Bernardo de Sá Mota e Mark G. Sourdoutovich pelas inúmeras conversas sobre minhas não raras dúvidas e pelas discussões de alto nível que tornaram esses anos breves.

Finalmente, não poderia de deixar de agradecer também aos meus pais e à minha namorada Cristiane Ueno. O seu incentivo e apoio em todos os momentos são de valor inestimável.
1 Introduction

The assumption of stationarity required in some statistical models makes them inadequate to analyze many series of economic data. This fact caused a great deal of research to be conducted in order to relax such an annoying assumption.

A simple way to reach this goal consists in modelling a series as a process which is stationary around a deterministic trend. This allows one to model series whose mean is clearly nonstationary and the tools needed to deal with such processes can be easily extended from the techniques employed for stationary processes. The development of these models, along with more recent advances in dealing with heterogeneously distributed errors by White (1980) and White & Domowitz (1984), rendered applied workers generally satisfied with the tools available for data analysis.

However, the study reported in Nelson & Plosser (1982) greatly influenced to change this situation. Applying the test developed by Dickey & Fuller (1979) to macroeconomic time series, they concluded that some of the series analyzed would be better modelled as integrated processes. The effects of spuriously detrending an integrated time series may be quite harmful, including a bias to accept the hypothesis of a nonzero trend coefficient in a driftless random walk (Nelson & Kang (1981)) and the rise of spurious correlations in the detrended series.

Besides providing guidance to model selection, the ability to detect the presence of a unit root may be relevant to test economic theories. In an integrated process, shocks are persistent and mean reversion does not take place. Therefore, hypotheses such as the Purchasing Power Parity in which reversion to a long run equilibrium is assumed can be tested by unit root tests as in Flöres, Jorion, Preumont & Szafarz (1999).

More recently, Granger (1981) introduced the concept of cointegration, in which two or more integrated series exhibit a common stochastic trend. This is of great importance to economic theory, since it allows one to detect long run equilibrium relationships among nonstationary series. Since cointegration analysis presumes integrated series, it is necessary to somehow be able to test the presence of a unit root in a time series.

The development of new methods to test the existence of a unit root became necessary, since the standard test statistics usually employed no longer apply. One of the first and most popular methods for testing this hypothesis was proposed, in its original version, in Dickey & Fuller (1979). In that paper, the asymptotic distribution of two different test statistics
are derived when a driftless random walk is the true data generating process (DGP). It is shown that, under such circumstances, the inclusion of a constant and a trend has influence on the distribution of both test statistics.

An important extension of the Dickey & Fuller test is made in Phillips (1987a) and Phillips & Perron (1988). In those papers milder assumptions about the shocks are made and the asymptotic distributions determined. One important result is the proof that, when a unit root is present, the OLS estimate of the autoregressive parameter is consistent regardless of serial correlation in the shocks. Adjustments in the test statistics are made to account for the error serial correlation and the resulting statistics have the same asymptotic distribution of those presented in Dickey & Fuller (1979).

However popular the DF and PP tests proposed by Dickey & Fuller (1979) and Phillips & Perron (1988) respectively are, they do offer some inconveniences. Firstly, both tests suffer from low power when dealing with alternatives close to the null. Although one may argue that any test proposed will have poor performance for some alternative close enough to the null, this does not mean that no effort should be employed in order to improve on the available ones. Some economic series appear to have long persistence and being able to reliably test them for unit roots would be rather desirable. Hansen (1995) suggested exploring the correlations between the series being tested and a series whose order of integration is known as a way of increasing power. It is important to notice that, despite using more than one series, this test will only deal with the existence of a unit root in one of them, which makes evident another problem suffered by the DF and PP tests: its inability to deal with multivariate series.

Phillips & Durlauf (1986) provided some useful results concerning a unit root test in a multivariate environment. The consistency of the OLS estimator under a unit root in each series is established and the consistency of the covariance matrix estimator follows. These results are derived under mild restrictions on the serial correlation of shocks. The test then proposed can be thought of as a multivariate version of the Phillips & Perron (1988) test.

In Abuaf & Jorion (1990), a multivariate version of a unit root test is used to test the PPP hypothesis. In the test proposed, a restriction of common dynamic is imposed and the estimation of parameters is made by minimizing a quadratic form of the residuals. The weighting matrix used is the inverse of the long run covariance matrix between the series. The authors provide Monte Carlo results showing that there is a huge gain in power in relation to the DF univariate test. Such improvement in performance is attributed partly
to the use of information contained in the covariance matrix and partly to the restriction of common dynamics.

The FPS test presented in Flôres, Preumont & Szafarz (1996) can be thought of as a generalization of the Abuaf & Jorion (1990) test, since the restriction of common dynamic is relaxed. In that paper, asymptotic results are derived which allow a variety of tests to be performed. The estimation procedure proposed presents two main differences from that proposed in Phillips & Durlauf (1986): the estimation of parameters is made by using Zellner’s SURE method instead of using the OLS method and the coefficients matrix of the VAR assumed for the data generated process is imposed to be diagonal.

In this essay, interest lies in obtaining results allowing one to determine the gain in power of the FPS test over the PD test and what fraction of it is due to the use of the restriction that the matrix of coefficients is diagonal for a certain class of alternative hypothesis. In order to do so, the results related to testing for the presence of unit roots presented in Phillips & Durlauf (1986) will be adapted to the case where a restriction on the matrix of coefficients is imposed. This allows us to determine how much power is added because of this restriction. The remaining difference in power between the FPS test and the restricted PD test shall be attributed to the use of a possibly more efficient weighting matrix in the estimations.

Since our interest is in the asymptotic power, we must be able to determine the asymptotic behavior of the test statistics under the class of alternative hypotheses being considered. In order to do so, we adopt the class of alternatives presented in Phillips (1988) and adapt those results to both the PD and FPS tests as well as to a new test proposed. The class of alternative hypotheses considered in this paper are a sequence of drifting DGPs which converge to the null in a suitable rate, following the original ideas in Pitman (1949) and Neyman (1937).

The remainder of this essay goes as follows. Before section 3 describes each of the tests considered in detail and derives their asymptotic distributions under the null hypothesis, section 2 describes a convergence result which will be useful throughout this essay. Section 4 describes the set of alternative hypotheses for which the distributions of the test will be derived and determines such distributions. Finally, section 5 concludes.
2 Functional Limit Results

Before we discuss the unit root tests themselves, a review of the results derived in Phillips & Durlauf (1986) is presented. Those results consist of a multivariate generalization of the univariate results due to McLeish (1975a) used in Phillips (1987a) and will be much helpful in the following sections.

Consider a sequence of random \( p \)-dimensional vectors \( \{u_t\} \) defined on a probability space \((\Omega, \mathcal{B}, P)\) such that:

\[
E(u_t) = 0, \forall t
\]  

(1)

Define the vector of partial sums:

\[
S_t = \sum_{j=1}^{t} u_j
\]  

(2)

and the vector random sequence of functions on \([0, 1]\):

\[
X_T(r) = \frac{1}{\sqrt{T}} \Sigma^{-\frac{1}{2}} S_{[T r]}, \quad \frac{(j-1)T}{T} \leq r < \frac{jT}{T}, \quad T = 1, 2, \ldots
\]

\[
X_T(1) = \frac{1}{\sqrt{T}} \Sigma^{-\frac{1}{2}} S_T
\]  

(3)

where \( r \in [0, 1] \) and \( \Sigma^{\frac{1}{2}} \) is the symmetric positive definite square root of the positive definite matrix \( \Sigma = \lim_{T \to \infty} E(T^{-1}S_T S_T') \)

We now describe a result concerning the asymptotic behavior of the process \( X_T(r) \) which will be helpful when we turn our attention to the limiting distribution of our test statistics under different conditions. A set of assumptions on the temporal dependence of the process \( \{u_t\} \) is made and an appropriate metric chosen in order to establish such a result.

In order to determine the metric under which convergence will be established note that:

\[
X_T(\bullet) \in D[0, 1]^p = D[0, 1] \times D[0, 1] \times \cdots \times D[0, 1]
\]  

(4)

where \( D[0, 1] \) is the space of all real valued functions from \([0, 1]\) to the real line that are right continuous at each point of \([0, 1]\) and possess finite left limits. The Skorohod metric \( d \) is chosen in order to turn the metric space \((D[0, 1], d)\) into a complete separable metric space. For the \( n \)-dimensional product space, the metric chosen is:

\[
d'(x, y) = \max_i \{d(x_i, y_i) : x_i, y_i \in D[0, 1], i = 1, \ldots, n\}
\]  

(5)
Once an appropriate metric for the space where $X_T(r)$ lies is chosen, a way to measure the temporal dependence of the process $\{u_t\}$ must be determined before any necessary restriction can be made. We now describe the measure of dependence employed in Phillips & Durlauf (1986). For the $\sigma$-algebras $F, G \subseteq \mathcal{B}$, define:

$$
\varphi(F, G) = \sup_{\{F \in \mathcal{F}, G \in \mathcal{G} \mid P(F) > 0\}} |P(G|F) - P(G)|
$$

$$
\alpha(F, G) = \sup_{\{F \in \mathcal{F}, G \in \mathcal{G}\}} |P(G|F)P(F) - P(F, G)|
$$

(6)

Let $a < b$ and $F_a^b$ be the $\sigma$-algebra generated by $\{u_a, \ldots, u_b\}$ and $R_a^b$ be the $\sigma$-algebra generated by $\{S_b - S_{a-1}, \forall a \leq b\}$. The temporal dependence of $\{u_t\}$ is measured by:

$$
\varphi_m = \sup_n sup_{j \geq n+m} \varphi(F_1^n, R_j^{n+m})
$$

$$
\alpha_m = \sup_n sup_{j \geq n+m} \alpha(F_1^n, R_j^{n+m})
$$

(7)

We say that $\varphi_m$ (or $\alpha_m$) is of size $-p$ if $\varphi_m$ (or $\alpha_m$) = $O(m^{-p-\epsilon})$ for some $\epsilon > 0$ as $m \uparrow \infty$.

Having described the metric for which convergence will be derived and the measures of temporal dependence which will be used, we are now ready to state a result concerning the asymptotic behavior of $X_T(r)$.

**Theorem 1** Let $\{u_t\}_{t=1}^{\infty}$ be a weakly stationary sequence of random $p$-dimensional vectors with $E(u_t) = 0$, $\forall t$. If:

1. $E|u_t|^\beta < \infty$, $\forall i = 1, \ldots, p$ for some $2 \leq \beta < \infty$;

2. either $\sum_{n=1}^{\infty} \varphi_n^{1-\frac{1}{\beta}} < \infty$ OR $\beta > 2$ and $\sum_{n=1}^{\infty} \alpha_n^{1-\frac{2}{\beta}} < \infty$

Then:

$$
\lim_{T \to \infty} E(T^{-1}S_TS_T') = \Sigma = E(u_1u_1') + \sum_{k=2}^{\infty} [E(u_1u_k') + E(u_ku_1')]
$$

If, furthermore, $\Sigma$ is positive definite, then

$$
X_T(r) \Rightarrow W(r)
$$

as $T \uparrow \infty$, where $\{W(r)\}$ is the standard $p$-dimensional Brownian Motion.

In Phillips & Durlauf (1986), conditions that assure the convergence of $X_T$ to $W$ when \( \{u_t\} \) is not identically distributed are also obtained. However, for that convergence to hold we must have $\lim_{T \to \infty} E(T^{-1}S_T S_T') = \Sigma$ as a hypothesis. Under stationarity and the conditions of Theorem 1, not only $X_T$ converges to $W$, but also $T^{-1}S_T S_T'$ converges to $\Sigma$. This will be important as consistent estimation of the long run covariance matrix $\Sigma$ will be involved in the test procedures under analysis.

3 Description of the Testing Procedures

In this section, the three tests will be described in detail. We begin by describing the common features and assumptions of the procedures and briefly discussing the relationships between them. The details of each procedure, including the test statistics to be used and their distributions under $H_0$, are then described. The distributions of the test statistics under $H_1$ are given in section 4.

All testing procedures considered in this paper are based on the following model:

$$y_t = Ay_{t-1} + u_t$$

(8)

where

\(\{y_t\}_{t=0}^T\) is a sequence of $p$-dimensional random vectors which will be tested for the presence of unit roots;

$A$ is a $p \times p$ matrix of coefficients;

$\{u_t\}_{t=0}^T$ is a sequence of $p$-dimensional shocks.

Throughout this work, the assumption below will be made:

Assumption 1 The process \(\{u_t\}\) is given by:

$$u_t = \Phi(L)e_t$$

(9)

where:

1. $\Phi(L) = \sum_{k=0}^{\infty} \Phi_k L^k$
2. $\Phi(1)$ is a non singular matrix;

3. $E(\varepsilon_t) = 0, \forall t$;

4. $E(\varepsilon_t\varepsilon_t') = I_p, \forall t$;

5. $E(\varepsilon_t\varepsilon_{t-j}) = 0, \forall t, \forall j \neq 0$;

6. $\varepsilon_t$ has finite fourth moments;

7. All eigenvalues of $\Phi(L)$ lie inside the unit circle;

8. $\sum_{k=0}^{\infty} \Phi_k \Phi_k'$ is a positive definite matrix;

Under such conditions, it is possible to assure that $\{u_t\}$ is a weakly stationary process satisfying all conditions under which the results of Theorem 1 are established. The $p \times p$ matrix $\Sigma_0$ will denote the contemporaneous covariance matrix, while $\Sigma$ will be used to refer to the long run covariance matrix:

$$\Sigma_0 = E(u_t u_t') = \sum_{k=0}^{\infty} \Phi_k \Phi_k'$$

$$\Sigma = \Phi(1)\Phi(1)'$$

Under the assumptions made, $\Sigma$ is a positive definite matrix.

Besides the long run and contemporaneous covariance matrices, the autocovariance matrices will appear in some asymptotic distributions. We define $\Gamma_j$, the autocovariance matrix of $j$-th order, by:

$$\Gamma_j = \sum_{k=j}^{\infty} (\Phi_k \Phi_{k-j}'), \ j \geq 1$$

and we define:

$$\Sigma_1 = \sum_{j=1}^{\infty} \Gamma_j$$

It is easy to show that:

$$\Sigma = \Sigma_0 + \Sigma_1 + \Sigma_1'$$
The hypothesis we will be interested in is given by:

\[ H_0 : A = I_p \]
\[ H_1 : A \neq I_p \]  \hspace{1cm} (15)

where \( I_p \) is the \( p \)-dimensional identity matrix and the class of matrices \( A \) must be duly defined.

Our main goal consists in determining a method in order to compare the asymptotic power of the multivariate unit root test proposed in Flôres, Preumont & Szafarz (1996) – the FPS procedure – and that described in Phillips & Durlauf (1986) – the PD procedure. Two main differences exist between these two testing procedures:

- The FPS test imposes a restriction \textit{a priori} that \( A \) is diagonal, while in the PD procedure \( A \) is unrestricted;

- The FPS test estimates the matrix of coefficients \( A \) using the seemingly unrelated regressions technique, which can be considered – in their case – a least squares where the weighting matrix is given by \( \Sigma^{-1} \) whereas the PD test relies on OLS estimation;

Since we are also interested in providing a way to determine what fraction of any difference in power is due to each of these peculiarities in the FPS test, a procedure in between the FPS and the PD is proposed.

Such testing procedure will impose the restriction that the matrix of coefficients \( A \) is diagonal as in the FPS test, but its estimation, similarly to what happens in the PD test, will make use of OLS. This can be thought of as a restricted PD test and therefore shall be referred to as the RPD test.

One could be tempted to propose a second testing procedure in which no restriction on \( A \) is made but OLS estimation is replaced by SURE. Although such a test may seem to stand in between the FPS and PD tests, its results will be equivalent to those of the PD test. This happens because the set of regressors will be the same for all \( p \) equations. Under these circumstances, the SURE can be efficiently estimated by using OLS. It is worth noting that when one imposes that \( A \) is diagonal, the set of regressors for each equation will no longer be the same and therefore the use of \( \Sigma^{-1} \) as weighting matrix may lead to a more efficient estimate than that resulting from the use of OLS.
Before we proceed to the description of the details of each test, a result which will be very helpful to determine the consistency of some of the estimators proposed and the asymptotic distribution of some test statistics under $H_0$ is presented:

**Theorem 2** Suppose $\{y_t\}$ is given by (8) with $A = I_p$ and $\{u_t\}$ satisfies the assumptions in Theorem 1. Then:

1. $T^{-\frac{3}{2}} \sum_{t=1}^{T} y_t \Rightarrow \Sigma^\frac{1}{2} \int_{0}^{1} W(r) dr$

2. $T^{-2} \sum_{t=1}^{T} y_{t-1} y'_{t-1} \Rightarrow \Sigma^\frac{1}{2} \int_{0}^{1} W(r) W(r)' dr \Sigma^\frac{1}{2}$

3. $T^{-1} \sum_{t=1}^{T} (y_{t-1} u'_t + u_t y'_{t-1}) \Rightarrow \Sigma^\frac{1}{2} W(1) W(1)' \Sigma^\frac{1}{2} - \Sigma_0$

4. $T^{-1} \sum_{t=1}^{T} u'_t \Rightarrow \Sigma^\frac{1}{2} \int_{0}^{1} W(r) dW(r)' \Sigma^\frac{1}{2} + \Sigma_1$

where:

$$\Sigma_0 = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} E(u_t u'_t)$$

$$\Sigma_1 = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} \sum_{j=1}^{T-1} E(u_j u'_t)$$

$$\Sigma = \lim_{T \to \infty} E(T^{-1} S_T S'_T) = \Sigma_0 + \Sigma_1 + \Sigma'_1$$

When $\{u_t\}$ satisfies assumption 1, $\Sigma_0$, $\Sigma$ and $\Sigma_1$ are as defined in (10), (11) and (13) respectively.

**Proof.** See Phillips & Durlauf (1986).

### 3.1 The PD Unit Root Test

Define:

$$\mathcal{Y}_{-1} = \begin{bmatrix} y_0 & y_1 & \cdots & y_{T-1} \end{bmatrix}$$

$$\mathcal{Y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_T \end{bmatrix}$$

so that $\mathcal{Y}_{-1}$ and $\mathcal{Y}$ are both $p \times T$ matrices.
Using OLS, two different consistent estimators under $H_0$ for the matrix $A$ can be obtained (see Appendix A):

\begin{align}
\hat{A}_{PD1} &= (\mathcal{YY}'_1)(\mathcal{Y}'_1\mathcal{Y}'_1)^{-1} \\
\hat{A}_{PD2} &= (\mathcal{Y}'_1\mathcal{Y}')(\mathcal{Y}'_1\mathcal{Y}'_1)^{-1}
\end{align}

Phillips & Durlauf (1986) prove that, under $H_0$, $\hat{A}_{PD1} \xrightarrow{p} I_p$ and $\hat{A}_{PD2} \xrightarrow{p} I_p$. A third estimator can be obtained by combining these two:

\begin{align}
\hat{A}_{PD} &= \frac{1}{2}(\hat{A}_{PD1} + \hat{A}_{PD2}) \\
&= \frac{1}{2} ((\mathcal{Y}'_1\mathcal{Y}') + (\mathcal{YY}'_1)) (\mathcal{Y}'_1\mathcal{Y}'_1)^{-1}
\end{align}

It follows immediately from the consistency of $\hat{A}_{PD1}$ and $\hat{A}_{PD2}$ and the symmetry of $A$ under $H_0$ that $\hat{A}_{PD} \xrightarrow{p} I_p$ in that case. A result proven in the same paper concerns the consistency of the estimator of the contemporaneous covariance matrix under the same situation:

\begin{align}
\hat{\Sigma}_{0PD} &= T^{-1} \mathcal{Y} \left( I_T - \mathcal{Y}'_1(\mathcal{Y}'_1\mathcal{Y}'_1)^{-1}\mathcal{Y}'_1 \right) \mathcal{Y}'
\end{align}

Before discussing the test statistics, we define an estimator for the long run covariance matrix $\Sigma$. The Newey and West (1987) estimator for $\Sigma$ will be used. It is given by:

\begin{align}
\hat{\Sigma}_{PD} &= T^{-1} \sum_{t=1}^{T} \tilde{u}_{PDt}\tilde{u}'_{PDt} \\
&+ T^{-1} \sum_{\tau=1}^{T} v_{\tau}\sum_{t=\tau+1}^{T} (\tilde{u}_{PDt}\tilde{u}'_{PDt-\tau} + \tilde{u}_{PDt-\tau}\tilde{u}'_{PDt})
\end{align}

with:

\begin{align}
v_{\tau} &= 1 - \frac{\tau}{T+1} \\
\tilde{u}_{PDt} &= y_t - \hat{A}_{PD}y_{t-1}
\end{align}

The next result assures the consistency of $\hat{\Sigma}_{PD}$

**Proposition 1** Suppose $\{y_t\}$ is given by (8) with $A = I_p$ and $\{u_t\}$ satisfies the assumptions in theorem 1. Suppose further that:
1. \( \sup_{t, i} E|u_{it}|^{2\beta} < \infty \) for some \( \beta > 2 \);

2. \( \lim_{T \to \infty} l_T = +\infty \);

3. \( l = o(T^{\frac{1}{4}}) \);

Then:

\[ \tilde{\Sigma}_{PD} \xrightarrow{p} \Sigma \]

**Proof.** Let

\[ \tilde{S}_T = T^{-1} \sum_{t=1}^{T} u_{it}u_{it}^\prime + T^{-1} \sum_{\tau=1}^{l} \sum_{t=\tau+1}^{T} (u_{it-\tau} + u_{t-\tau}u_{i}) \]

If \( l \) and \( u_t \) satisfy the conditions above, Theorem 3.6 in Phillips & Durlauf (1986) guarantees that \( \tilde{S}_T \xrightarrow{p} \Sigma \) as \( T \uparrow \infty \). Choosing the same sequence for \( l \), we have that \( v_{-t} \to 1 \) as \( T \uparrow \infty \) and therefore \( \tilde{\Sigma}_{PD} - \tilde{S}_T \xrightarrow{P} 0 \). Since \( \tilde{A}_{PD} \) is consistent, \( \tilde{u}_{PDt} \) converges in probability to \( u_{PDt}, \forall t \) and the proof is complete.

In Phillips & Durlauf (1986) the following test statistic, based on a Wald test, is proposed to test \( H_0 \):

\[
F_{0PD} = \text{vec}(\tilde{A}_{PD} - I_p)' \left[ \tilde{\Sigma}_{0PD} \otimes (Y_{-i-1})^{-1} \right]^{-1} \text{vec}(\tilde{A}_{PD} - I_p) = \text{tr} \left[ (\tilde{A}_{PD} - I_p)\tilde{\Sigma}_{0PD}^{-1}(\tilde{A}_{PD} - I_p)(Y_{-i-1}) \right] \quad (23)
\]

If instead of the weight matrix used in (23), we use as weight matrix an estimate of the inverse of \( \left[ \Sigma \otimes (Y_{-i-1})^{-1} \right] \), the test statistics become:

\[
F_{PD} = \text{vec}(\tilde{A}_{PD} - I_p)' \left[ \tilde{\Sigma}_{PD} \otimes (Y_{-i-1})^{-1} \right]^{-1} \text{vec}(\tilde{A}_{PD} - I_p) = \text{tr} \left[ (\tilde{A}_{PD} - I_p)\tilde{\Sigma}_{PD}^{-1}(\tilde{A}_{PD} - I_p)(Y_{-i-1}) \right] \quad (24)
\]

Having defined the test statistics used in this test procedure, we now turn our attention to their distribution under \( H_0 \).
Proposition 2. Suppose that conditions in Proposition 1 are satisfied and $F_{0PD}$ and $F_{PD}$ are defined as in (23) and (24) respectively. Then:

$$F_{0PD} \Rightarrow \frac{1}{4} \text{tr} \left\{ \left[ W(1)W(1)' - \Sigma^{-rac{1}{2}}\Sigma_0\Sigma^{-rac{1}{2}} \right] \Sigma^{rac{1}{2}}\Sigma_0^{-1}\Sigma^{rac{1}{2}} \left[ W(1)W(1)' - \Sigma^{-rac{1}{2}}\Sigma_0\Sigma^{-rac{1}{2}} \right] \times \right\} \left( \int_0^1 W(r)W(r)'dr \right)^{-1} \right\}$$  \hspace{1cm} (25)

$$F_{PD} \Rightarrow \frac{1}{4} \text{tr} \left\{ \left[ W(1)W(1)' - \Sigma^{-rac{1}{2}}\Sigma_0\Sigma^{-rac{1}{2}} \right] \left[ W(1)W(1)' - \Sigma^{-rac{1}{2}}\Sigma_0\Sigma^{-rac{1}{2}} \right] \times \right\} \left( \int_0^1 W(r)W(r)'dr \right)^{-1} \right\}$$  \hspace{1cm} (26)

Proof. For the first convergence, see Phillips & Durlauf (1986). For the second one, note that:

$$F_{PD} = \text{tr} \left( \left( \frac{T}{2} \sum_{t=1}^T y_{t-1}u_t' + u_t'y_{t-1}' \right) \left( \frac{T}{2} \sum_{t=1}^T y_{t-1}y_{t-1}' \right)^{-1} \tilde{\Sigma}_{PD} \left( \frac{T}{2} \sum_{t=1}^T y_{t-1}u_t' + u_t'y_{t-1}' \right) \times \right)$$

where the second equality comes from the fact that $(\tilde{A}_{PD} - I_p)\tilde{\Sigma}_{PD}^{-1}(\tilde{A}_{PD} - I_p)(Y_{-1}Y_{-1}')$ is symmetric. Using the results in theorem 2, it follows that:

$$F_{PD} \Rightarrow \text{tr} \left[ \frac{1}{4} \left( \Sigma^{rac{1}{2}}W(1)W(1)\Sigma^{rac{1}{2}} - \Sigma_0 \right) \Sigma^{-1} \left( \Sigma^{rac{1}{2}}W(1)W(1)\Sigma^{rac{1}{2}} - \Sigma_0 \right) \times \right] \left( \Sigma^{rac{1}{2}} \int_0^1 W(r)W(r)'dr \Sigma^{rac{1}{2}} \right)^{-1}$$

It is possible to make corrections in both test statistics in order to obtain a distribution under $H_0$ free of nuisance parameters. The test statistics obtained by performing such
corrections are:

\[ F_{0SPD} = F_{0PD} - \frac{1}{4} \text{tr} \left\{ \left[ T^{-2} y_T y_T (\tilde{\Sigma}_{0PD}^{-1} - \tilde{\Sigma}_{PD}^{-1}) y_T y_T + (\Sigma_{0PD} - \Sigma_{PD}) \right] \times (T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1}) \right\} \ \ (27) \]

\[ F_{SPD} = F_{PD} - \frac{1}{4} \text{tr} \left\{ \left[ T^{-1} y_T y_T \tilde{\Sigma}_{PD}^{-1} (\tilde{\Sigma}_{PD} - \Sigma_{0PD}) + (\tilde{\Sigma}_{PD} - \Sigma_{0PD}) \tilde{\Sigma}_{PD}^{-1} T^{-1} y_T y_T + (\tilde{\Sigma}_{PD} - \Sigma_{0PD}) \tilde{\Sigma}_{PD}^{-1} (\tilde{\Sigma}_{PD} - \Sigma_{0PD}) - 2(\tilde{\Sigma}_{PD} - \Sigma_{0PD}) \right] (T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1}) \right\} \ \ (28) \]

We now turn our attention to determining the distribution under \( H_0 \) of these test statistics:

**Proposition 3** Suppose that conditions in Theorem 1 are satisfied and \( F_{0SPD} \) and \( F_{SPD} \) are as defined in (27) and (28). Then:

\[ F_{0SPD} \Rightarrow \frac{1}{4} \text{tr} \left\{ [W(1)W(1)' - I_p] [W(1)W(1)' - I_p] \left[ \int_0^1 W(r)W(r)' \text{d}r \right]^{-1} \right\} \ \ (29) \]

\[ F_{SPD} \Rightarrow \frac{1}{4} \text{tr} \left\{ [W(1)W(1)' - I_p] [W(1)W(1)' - I_p] \left[ \int_0^1 W(r)W(r)' \text{d}r \right]^{-1} \right\} \ \ (30) \]

**Proof.** For the first convergence, define:

\[ C_1 = \text{tr} \left[ T^{-2} y_T y_T (\tilde{\Sigma}_{0PD}^{-1} - \tilde{\Sigma}_{PD}^{-1}) y_T y_T (T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1}) \right] \]

\[ C_2 = \text{tr} \left[ (\tilde{\Sigma}_{0PD} - \Sigma_{PD}) (T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1}) \right] \]

It follows that:

\[ C_1 \Rightarrow \text{tr} \left[ (W(1)W(1)' \Sigma^{-\frac{1}{2}} (\Sigma_0^{-1} - \Sigma^{-1}) \Sigma^{-\frac{1}{2}} W(1)W(1)') \left( \int_0^1 W(r)W(r)' \text{d}r \right)^{-1} \right] \]

\[ C_2 \Rightarrow \text{tr} \left[ (\Sigma^{-\frac{1}{2}} (\Sigma_0 - \Sigma) \Sigma^{-\frac{1}{2}}) \left( \int_0^1 W(r)W(r)' \text{d}r \right)^{-1} \right] \]

In Appendix B it is proved that:

\[ (W(1)W(1)' - \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}}) (\Sigma^{-\frac{1}{2}} \Sigma_0^{-1} \Sigma^{-\frac{1}{2}}) (W(1)W(1)' - \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}}) = (W(1)W(1)' - I_p) (W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}} (\Sigma_0 - \Sigma) \Sigma^{-\frac{1}{2}}) + (W(1)W(1)' \Sigma^{-\frac{1}{2}} (\Sigma_0^{-1} - \Sigma^{-1}) \Sigma^{-\frac{1}{2}} W(1)W(1)') \]
Therefore:

\[ F_{0PD} \Rightarrow \frac{1}{4} \text{tr} \left\{ \left[ (W(1)W(1)' - I_p) (W(1)W(1)' - I_p) \right] \left( \Sigma^{-\frac{1}{2}} (\Sigma_0 - \Sigma) \Sigma^{-\frac{1}{2}} \right) \right. \]

\[ + \left. \left( W(1)W(1)' \Sigma^{-\frac{1}{2}} (\Sigma_0^{-1} - \Sigma^{-1}) \Sigma^{-\frac{1}{2}} W(1)W(1)' \right) \right\} \left[ \int_0^1 W(r)W(r)'dr \right]^{-1} \]

It follows that:

\[ F_{0SPD} = F_{0PD} - \frac{1}{4} C_1 - \frac{1}{4} C_2 \]

\[ \Rightarrow \frac{1}{4} \text{tr} \left\{ \left[ (W(1)W(1)' - I_p) (W(1)W(1)' - I_p) \right] \left[ \int_0^1 W(r)W(r)'dr \right]^{-1} \right\} \]

For the second convergence, define:

\[ C_1 = \text{tr} \left[ \left( T^{-1}y_Ty_T^T \Sigma_{PD}^{-1} (\tilde{\Sigma}_{PD} - \tilde{\Sigma}_{0PD}) \right) \left( T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1} \right) \right] \]

\[ = \text{tr} \left[ \left( (\tilde{\Sigma}_{PD} - \tilde{\Sigma}_{0PD}) \Sigma_{PD}^{-1} (\tilde{T}^{-1}y_Ty_T^T) \right) \left( T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1} \right) \right] \]

\[ C_2 = \text{tr} \left[ \left( (\tilde{\Sigma}_{PD} - \tilde{\Sigma}_{0PD}) \Sigma_{PD}^{-1} (\tilde{T}^{-1}y_Ty_T^T) \right) \left( T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1} \right) \right] \]

\[ C_3 = \text{tr} \left[ \left( (\tilde{\Sigma}_{PD} - \tilde{\Sigma}_{0PD}) \right) \left( T^{-2} \mathcal{Y}_{-1} \mathcal{Y}_{-1}^{-1} \right) \right] \]

It follows that:

\[ C_1 \Rightarrow \text{tr} \left[ \left( \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} W(1)W(1)' \right) \left( \int_0^1 W(r)W(r)'dr \right)^{-1} \right] \]

\[ \Rightarrow \text{tr} \left[ \left( W(1)W(1)' \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \right) \left( \int_0^1 W(r)W(r)'dr \right)^{-1} \right] \]

\[ C_2 \Rightarrow \text{tr} \left[ \left( \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \right) \left( \int_0^1 W(r)W(r)'dr \right)^{-1} \right] \]

\[ C_3 \Rightarrow \text{tr} \left[ \left( \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \right) \left( \int_0^1 W(r)W(r)'dr \right)^{-1} \right] \]

Besides:

\[ \left[ W(1)W(1)' - \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}} \right] \left[ W(1)W(1)' - \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}} \right] = \]

\[ (W(1)W(1)' - I_p) + \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \times \]

\[ (W(1)W(1)' - I_p) + \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} = \]

\[ ((W(1)W(1)' - I_p)(W(1)W(1)' - I_p) \]

14
\[ + \left( W(1) W(1)' - I_p \right) (\Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}}) + \\
\left( \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \right) (W(1) W(1)' - I_p) + \\
\left( \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \right) (\Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}}) = \\
\left[ (W(1) W(1)' - I_p) (W(1) W(1)' - I_p) + \\
W(1) W(1)' \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} + \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} W(1) W(1)' + \\
2 (\Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}}) + (\Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}}) (\Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}}) \right] \\
\]

Therefore:

\[ F_{PD} \Rightarrow \frac{1}{4} tr \left\{ \left[ (W(1) W(1)' - I_p) (W(1) W(1)' - I_p) \\
+ \left( W(1) W(1)' \Sigma^{-\frac{1}{2}} \Sigma^{-1} (\Sigma_0 - \Sigma) \Sigma^{-\frac{1}{2}} \right) \\
+ \left( \Sigma^{-\frac{1}{2}} (\Sigma_0 - \Sigma) \Sigma^{-1} \Sigma^{-\frac{1}{2}} W(1) W(1)' \right) \\
+ \left( \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}} \right) + \\
- 2 (\Sigma^{-\frac{1}{2}} (\Sigma - \Sigma_0) \Sigma^{-\frac{1}{2}}) \left[ \int_0^1 W(r) W(r)' dr \right]^{-1} \right\} \]

It results that:

\[ F_{SPD} = F_{PD} - 2 C_1 - C_2 + 2 C_3 \]

\[ F_{SPD} \Rightarrow \frac{1}{4} tr \left\{ \left[ (W(1) W(1)' - I_p) (W(1) W(1)' - I_p) \right] \left[ \int_0^1 W(r) W(r)' dr \right]^{-1} \right\} \]

3.2 The RPD Unit Root Test

As mentioned above, in the RPD testing procedure, we will impose a restriction that the matrix of coefficients \( A \) is diagonal. By measuring the performance of such a procedure, we may get a picture of how much of any difference in power between the FPS and the PD procedures is due to the restriction that the matrix of coefficients is diagonal used in the former. Taking the restriction into account, we are left with the following model:

\[ y_t = A_R y_{t-1} + u_t \]
\[
\begin{bmatrix}
y_{1t} \\
y_{2t} \\
\vdots \\
y_{pt}
\end{bmatrix} =
\begin{bmatrix}
a_1 & 0 & \cdots & 0 \\
0 & a_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & a_p
\end{bmatrix}
\begin{bmatrix}
y_{1t-1} \\
y_{2t-1} \\
\vdots \\
y_{pt-1}
\end{bmatrix}
+ 
\begin{bmatrix}
u_{1t} \\
u_{2t} \\
\vdots \\
u_{pt}
\end{bmatrix}
\tag{31}
\]

Defining:

\[\alpha = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \end{bmatrix}', \text{ a } p - \text{dimensional vector} \tag{32}\]

\[\iota_p = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}', \text{ a } p - \text{dimensional vector} \tag{33}\]

\[\phi_t = \text{diag}(y_t), \text{ a } p \times p \text{ matrix} \tag{34}\]

\[\mu_t = \text{diag}(u_t), \text{ a } p \times p \text{ matrix} \tag{35}\]

\[\Phi_{-1} = \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{T-1} \end{bmatrix}', \text{ a } Tp \times p \text{ matrix} \tag{36}\]

\[\Phi = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_T \end{bmatrix}', \text{ a } Tp \times p \text{ matrix} \tag{37}\]

the model becomes:

\[y_t = \phi_{t-1} \alpha + u_t \tag{38}\]

Using the notation just introduced, it is possible to rewrite the OLS estimate of matrix \(A\) for the restricted model in a way that will allow us to derive some results in a more general way, which will be useful when dealing with both the RPD and the FPS test. The RPD estimate \(\hat{A}_{RPD}\) of \(A\) is the diagonal matrix whose diagonal is defined by \(\hat{\alpha}_{RPD}\) below:

\[
\hat{\alpha}_{RPD} = \left( \sum_{t=1}^{T} \phi_{t-1}' \phi_{t-1} \right)^{-1} \left( \sum_{t=1}^{T} \phi_{t-1}' y_t \right) \\
= \left( \sum_{t=1}^{T} \phi_{t-1}' \phi_{t-1} \right)^{-1} \left( \sum_{t=1}^{T} \phi_{t-1}' \phi_{t-1} \phi_{t-1}' \phi_{t-1} y_{t-1} \right) \\
= \left( \Phi_{-1}' \Phi_{-1} \right)^{-1} \left( \Phi_{-1}' \Phi \right) \iota_p \tag{39}\]

As in the PD test, estimates of the contemporaneous and long run covariance matrices will be needed. For the long run covariance matrix, the Newey-West estimator will be once more employed. The estimates of these matrices are given by:

\[
\tilde{\Sigma}_{0RPD} = T^{-1} \sum_{t=1}^{T} \tilde{u}_t \tilde{u}_t' \tag{40}\]

16
\[ \hat{\Sigma}_{RDP} = T^{-1} \sum_{t=1}^{T} \hat{u}_t \hat{u}_t' + T^{-1} \sum_{\tau=1}^{l} v_{\tau l} \sum_{t=\tau+1}^{T} (\hat{u}_t \hat{u}_{t-\tau} + \hat{u}_{t-\tau} \hat{u}_t) \]  

where

\[ \hat{u}_t = y_t - \tilde{A}_{RDP} y_{t-1} \]
\[ v_{\tau l} = 1 - \frac{\tau}{l+1} \]

The results presented in Theorem 2 may be easily adapted to deal with the asymptotics related to the RPD and FPS procedure. The following proposition consists in a generalization of a result presented in Flôres, Preumont & Szafarz (1996) which will allow us to derive the asymptotic distributions to both tests.

**Proposition 4** Let \( A_R = I_p \) and \( y_t, \phi_t \) and \( \mu_t \) be given as in (31), (34) and (35) respectively and \( \Theta \) be a \( p \times p \) matrix. Assume that \( u_t \) satisfies the conditions listed in Assumption 1. Under these conditions, we have as \( T \uparrow \infty \):

1. \( T^{-2} \left( \sum_{t=1}^{T} \phi_{t-1}' \Theta \phi_{t-1} \right) \Rightarrow \Theta \odot \left( \Sigma_{\frac{1}{2}} \int_0^1 W(t)W(t)'dt \Sigma_{\frac{1}{2}}^{-1} \right) \)
2. \( T^{-1} \left( \sum_{t=1}^{T} \phi_{t-1}' \Theta \mu_{t} \right) \Rightarrow \Theta \odot \left( \Sigma_{\frac{1}{2}} \int_0^1 W(t)dW(t)' \Sigma_{\frac{1}{2}}^{-1} + \Sigma_{1} \right) \)

where \( \Sigma_{\frac{1}{2}} \) is as defined in Theorem 2 and \( \odot \) denotes the Hadamard product.

**Proof.** From (34) and (35), it follows that:

\[ T^{-2} \sum_{t=1}^{T} \phi_{t-1}' \Theta \phi_{t-1} = \Theta \odot T^{-2} \sum_{t=1}^{T} y_{t-1} y_{t-1}' \]
\[ T^{-1} \sum_{t=1}^{T} \phi_{t-1}' \Theta \mu_{t} = \Theta \odot T^{-1} \sum_{t=1}^{T} y_{t-1} u_{t}' \]

Combining these and Theorem 2, the results follow from the continuous mapping theorem. 

These results are helpful in establishing the consistency of \( \hat{A}_{RDP}, \hat{\Sigma}_0 \) and \( \hat{\Sigma}_{RDP} \) which is made in the next result by adapting the results from Phillips & Durlauf (1986) and Proposition 4 above to the case at hand:

17
Proposition 5 Suppose \( \{y_t\} \) is given by (38) with \( \alpha = \tau_p \) and \( \{u_t\} \) satisfies the assumptions in Theorem 1. Suppose further that:

1. \( \sup_{i,t} E|u_{it}|^{2\beta} < \infty \) for some \( \beta > 2 \);
2. \( \lim_{T \to \infty} l_T = +\infty \);
3. \( l_T = o(T^{1/2}) \);

Then:

1. \( \tilde{\alpha}_{RPD} \xrightarrow{p} \tau_p \)
2. \( \tilde{\Sigma}_0 \xrightarrow{p} \Sigma_0 \)
3. \( \tilde{\Sigma}_{RPD} \xrightarrow{p} \Sigma \)

Proof. Consistency of \( \tilde{\alpha}_{RPD} \) under \( H_0 \) follows from the fact that in such conditions:

\[
T(\tilde{\alpha}_{RPD} - \tau_p) = \left( T^{-2} \sum_{t=1}^{T} \phi_{t-1} \phi_{t-1}' \right)^{-1} \left( T^{-1} \sum_{t=1}^{T} \phi_{t-1} \mu_t \right) \tau_p
\]

\[
T(\tilde{\alpha}_{RPD} - \tau_p) \Rightarrow \left( I_p \odot \Sigma_0^2 \left( \int_0^1 W(r)W(r)'dr \Sigma_0^2 \right) \right)^{-1} \times \left( I_p \odot \Sigma_0^2 \left( \int_0^1 W(r)dW(r)' \Sigma_0^2 \right) \right) \tau_p
\]

Let \( Z \) be a random vector such that \( Z = \left( I_p \odot \Sigma_0^2 \int_0^1 W(r)W(r)'dr \Sigma_0^2 \right)^{-1} \times \left( I_p \odot \Sigma_0^2 \int_0^1 W(r)dW(r)' \Sigma_0^2 \right) \tau_p \). Since \( T(\tilde{\alpha}_{RPD} - \tau_p) \Rightarrow Z \), \( T(\tilde{\alpha}_{RPD} - \tau_p) \) is \( O_p(1) \) (See White (1999), Lemma 4.5). It follows that \( (\tilde{\alpha}_{RPD} - \tau_p) = \frac{1}{T} O_p(1) \) and since \( \frac{1}{T} \) is \( o_p(1) \), we have that \( (\tilde{\alpha}_{RPD} - \tau_p) \) is \( o_p(1) \) and therefore \( \tilde{\alpha}_{RPD} \) converges in probability to \( \tau_p \).

The consistency of \( \tilde{\Sigma}_{RPD} \) follows from the consistency of \( \tilde{\alpha}_{RPD} \) and an argument similar to that of Proposition 1.

In order to test test \( H_0 \), two test statistics based on the Wald test will be proposed. As in the PD test, the difference between the two relates to the weighting matrix used. The first one \( (F_{0RPD}) \) will be based on a weighting matrix using the contemporaneous covariance
matrix \( \Sigma_0 \), while the second one \((F_{RPD})\) makes use of the long run covariance matrix \( \Sigma \). They are defined by:

\[
F_{0RPD} = (\tilde{\alpha}_{RPD} - \iota_p)' \left[ (\Phi'_{-1} \Phi_{-1}) \left( I_T \otimes \tilde{\Sigma}_{0RPD} \right) \Phi_{-1} \right]^{-1} \left( \Phi'_{-1} \Phi_{-1} \right) \times (\tilde{\alpha}_{RPD} - \iota_p) \tag{42}
\]

\[
F_{RPD} = (\tilde{\alpha}_{RPD} - \iota_p)' \left[ (\Phi'_{-1} \Phi_{-1}) \left( I_T \otimes \tilde{\Sigma}_{RPD} \right) \Phi_{-1} \right]^{-1} \left( \Phi'_{-1} \Phi_{-1} \right) \times (\tilde{\alpha}_{RPD} - \iota_p) \tag{43}
\]

Finally, we turn our attention to the asymptotic distribution of the just defined test statistics under \( H_0 \):

**Proposition 6** Suppose \( \{y_t\} \) is given by (38) with \( \alpha = \iota_p \) and the conditions in Proposition 5 are observed. Then:

\[
F_{0RPD} \Rightarrow \iota_p' \left( I_p \otimes (\Sigma_0^\frac{1}{2} \int_0^1 dW(r)W(r)' \Sigma_1^\frac{1}{2} + \Sigma_1') \right) \times \left( \Sigma_0 \otimes \Sigma_0^\frac{1}{2} \int_0^1 W(r)W(r)' \Sigma_0^\frac{1}{2} \right)^{-1} \times \left( I_p \otimes (\Sigma_0^\frac{1}{2} \int_0^1 dW(r)dW(r)' \Sigma_0^\frac{1}{2} + \Sigma_1) \right) \iota_p \tag{44}
\]

\[
F_{RPD} \Rightarrow \iota_p' \left( I_p \otimes (\Sigma_0^\frac{1}{2} \int_0^1 dW(r)W(r)' \Sigma_1^\frac{1}{2} + \Sigma_1') \right) \times \left( \Sigma_0 \otimes \Sigma_0^\frac{1}{2} \int_0^1 W(r)W(r)' \Sigma_0^\frac{1}{2} \right)^{-1} \times \left( I_p \otimes (\Sigma_0^\frac{1}{2} \int_0^1 dW(r)dW(r)' \Sigma_0^\frac{1}{2} + \Sigma_1) \right) \iota_p \tag{45}
\]

**Proof.** Let \( \tilde{\Delta}_T \) be a sequence of \( p \times p \) positive definite matrices such that \( \tilde{\Delta}_T \rightarrow^p \Delta \). Consider the test statistics given by:

\[
F_{\Delta_T} = (\tilde{\alpha}_{RPD} - \iota_p)' \times \left[ (\Phi'_{-1} \Phi_{-1}) \left( I_T \otimes \tilde{\Delta}_T \right) \Phi_{-1} \right]^{-1} \left( \Phi'_{-1} \Phi_{-1} \right) \times (\tilde{\alpha}_{RPD} - \iota_p)
\]
\[
F_{\Delta_T} = \left[ \left( \sum_{t=1}^{T} \mu_t \phi_{t-1} \right) \left( \Phi_{-1}' \Phi_{-1} \right)^{-1} \times \left( \Phi_{-1}' \Phi_{-1} \right) \left( \Phi_{-1}' \left( I_T \otimes \tilde{\Delta}_T \right) \Phi_{-1} \right)^{-1} \left( \Phi_{-1}' \Phi_{-1} \right) \times \left( \Phi_{-1}' \Phi_{-1} \right)^{-1} \left( \sum_{t=1}^{T} \phi_{t-1} \mu_t \right) \right] \beta_p
\]

Using the results in proposition 4 it follows that:

\[
F_{\tilde{\Delta}_T} \Rightarrow \left[ \left( \sum_{t=1}^{T} u_t y_{t-1}' \right) \left( \tilde{\Delta}_T \otimes \sum_{t=1}^{T} y_{t-1} y_{t-1}' \right)^{-1} \left( \sum_{t=1}^{T} y_{t-1} u_t \right) \beta_p \right]
\]

The results proposed follow from setting \( \tilde{\Delta}_T = \tilde{\Sigma}_{0 RPD} \) and \( \tilde{\Delta}_T = \tilde{\Sigma}_{RPD} \).

3.3 The FPS Unit Root Test

As mentioned above, the FPS test bears some resemblance to the RPD test in that both make use of the restriction that \( A \) is diagonal. However, contrary to the RPD test, the FPS test uses a different weighting matrix based on the long run covariance matrix when minimizing the sum of squared residuals. We now describe this procedure in detail.

Under the conditions imposed on \( u_t \), \( \Sigma \) is a positive definite matrix and, as a consequence, it is possible to choose \( P \) positive definite (and symmetric), such that \( PP = \Sigma \). Premultiplying (38) by \( P^{-1} \), we get:

\[
P^{-1}y_t = P^{-1}\phi_{t-1}\alpha + P^{-1}u_t
\]

\[
y_t^* = \phi_{t-1}^*\alpha + u_t^*
\]

It is easy to notice that (46) is very similar to (38). The restrictions regarding \( u_t \) ensure that \( u_t^* \) is a weakly stationary process satisfying the assumptions made in Theorem 1.
As an estimate of $\alpha$, we shall use:

$$\tilde{\alpha}_{FPS} = \left[ \sum_{t=1}^{T} \phi'_{t-1} \phi^*_{t-1} \right]^{-1} \left[ \sum_{t=1}^{T} \phi'_{t-1} \phi^*_{t-1} \right] t_p$$

$$= \left[ \sum_{t=1}^{T} \phi'_{t-1} \Sigma^{-1} \phi^*_{t-1} \right]^{-1} \left[ \sum_{t=1}^{T} \phi'_{t-1} \Sigma^{-1} \phi^*_{t-1} \right] t_p$$

(47)

One point worth noting is that the estimator $\tilde{\alpha}_{FPS}$ as presented in (47) is unfeasible. However, a consistent estimate $\tilde{P}$ can be obtained from the spectral decomposition of any consistent estimate $\tilde{\Sigma}$ of the long run covariance matrix. We could for instance use the estimate given by $\tilde{\Sigma}_{RPD}$. Actually, the estimate $\tilde{\Sigma}_{PD}$ could also be used, since it is also consistent under $H_0$.

As in the other tests, estimates of the contemporaneous and long run covariance matrices are needed. They will be given by:

$$\tilde{\Sigma}_{0FPS} = T^{-1} \sum_{t=1}^{T} \tilde{u}_t \tilde{u}'_t$$

(48)

$$\tilde{\Sigma}_{FPS} = T^{-1} \sum_{t=1}^{T} \tilde{u}_t \tilde{u}'_t +$$

$$= T^{-1} \sum_{\tau=1}^{l} \nu_{\tau l} \sum_{t=\tau+1}^{T} (\tilde{u}_t \tilde{u}'_{t-\tau} + \tilde{u}_{t-\tau} \tilde{u}'_t)$$

(49)

where

$$\nu_{\tau l} = 1 - \frac{\tau}{l+1}$$

$$\tilde{u}_t = y_t - \tilde{A}_{FPS} y_{t-1}$$

$$= y_t - \phi_{t-1} \tilde{\alpha}_{FPS}$$

We must now present a result establishing the consistency of the estimates just defined. This is our next step.

**Proposition 7** Suppose $\{y_t\}$ is given by (38) with $\alpha = \iota_p$ and $\{u_t\}$ satisfies the assumptions in Theorem 1. Suppose further that:

1. $\sup_t E|u_{it}|^{2\beta} < \infty$ for some $\beta > 2$;
\[\lim_{T \to \infty} l = +\infty;\]
\[l = o(T^{\frac{1}{4}});\]

Then:

1. \(\hat{\alpha}_{FPS} \overset{p}{\to} \iota_p\)
2. \(\hat{\Sigma}_{0FPS} \overset{p}{\to} \Sigma_0\)
3. \(\hat{\Sigma}_{FPS} \overset{p}{\to} \Sigma\)

**Proof.** Consistency of \(\hat{\alpha}_{FPS}\) under \(H_0\) follows from the fact that in such conditions:

\[
T(\hat{\alpha}_{FPS} - \iota_p) = \left( T^{-2} \sum_{t=1}^{T} \phi_{t-1} \hat{\Sigma}_{FPS}^{-1} \phi'_{t-1} \right)^{-1} \left( T^{-1} \sum_{t=1}^{T} \phi_{t-1} \hat{\Sigma}_{FPS}^{-1} \mu'_t \right) \iota_p
\]

\[
T(\hat{\alpha}_{RPD} - \iota_p) = \left( \Sigma^{-1} \odot \sum_{t=1}^{T} W(r)W(r)'d\tau \Sigma_t^2 \right)^{-1} \times
\]
\[
\left( \Sigma^{-1} \odot (\Sigma^{\frac{1}{2}} \int_0^1 W(r)dW(r)'\Sigma^{\frac{1}{2}} + \Sigma_1) \right) \iota_p
\]

By an argument similar to that presented in Proposition 5, it is possible to show that this leads us to the consistency of \(\hat{\alpha}_{FPS}\).

The consistency of \(\hat{\Sigma}_{FPS}\) follows from the consistency of \(\hat{\alpha}_{RPD}\) and an argument similar to that of Proposition 1.

As in the previous cases, we shall use two different test statistics to test \(H_0\). They are similar to the ones presented above in that one of them is derived from a Wald test using the contemporaneous covariance matrix as weighting matrix, while the other one makes use of the long run covariance matrix. The two test statistics are presented below:

\[
F_{0FPS} = (\hat{\alpha}_{FPS} - \iota_p)' \left( \Phi^{-1} \hat{\Sigma}^{-1} \Phi^{-1} \right) \times
\]
\[
\left[ \Phi^{-1} \left( I_T \otimes \hat{\Sigma}_{0FPS} \hat{\Sigma}^{-1} \right) \Phi^{-1} \right]^{-1} \times
\]
\[
(\hat{\alpha}_{FPS} - \iota_p)
\]

\[
F_{FPS} = (\hat{\alpha}_{FPS} - \iota_p)' \times
\]
\[
\left\{ \left( \Phi^{-1} \hat{\Sigma}^{-1} \Phi^{-1} \right) \left[ \Phi^{-1} \Phi^{-1} \right]^{-1} \left( \Phi^{-1} \hat{\Sigma}^{-1} \Phi^{-1} \right) \right\} \times
\]
\[
(\hat{\alpha}_{FPS} - \iota_p)
\]
where \( \tilde{\Sigma} \) is any consistent estimate of \( \Sigma \) and \( \tilde{P} \) is a consistent estimate of \( \Sigma^{\frac{1}{2}} \).

Our next result concerns the asymptotic distributions under \( H_0 \) of the test statistics defined above.

**Proposition 8** Suppose \( \{y_t\} \) is given by (38) with \( \alpha = \nu_p \) and the conditions in proposition 7 are observed. Let \( F_{0\text{FPS}} \) and \( F_{\text{FPS}} \) be as defined as in (50) and (51) respectively. Then:

\[ F_{0\text{FPS}} \Rightarrow \nu_p \left[ \left( \Sigma^{-1} \otimes \left( \Sigma^{\frac{1}{2}} \int_0^1 dW(r)W(r)'\Sigma^{\frac{1}{2}} + \Sigma' \right) \right) \times \left( \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}} \otimes \left( \Sigma^{\frac{1}{2}} \int_0^1 W(r)W(r)'d\nu \Sigma^{\frac{1}{2}} \right) \right)^{-1} \times \left( \Sigma^{-1} \otimes \left( \Sigma^{\frac{1}{2}} \int_0^1 W(r)dW(r)'\Sigma^{\frac{1}{2}} + \Sigma_1 \right) \right) \right] \nu_p \] \tag{52}

\[ F_{\text{FPS}} \Rightarrow \nu_p \left[ \left( \Sigma^{-1} \otimes \left( \Sigma^{\frac{1}{2}} \int_0^1 dW(r)W(r)'\Sigma^{\frac{1}{2}} + \Sigma' \right) \right) \times \left( I_p \otimes \left( \Sigma^{\frac{1}{2}} \int_0^1 W(r)dW(r)'d\nu \Sigma^{\frac{1}{2}} \right) \right)^{-1} \times \left( \Sigma^{-1} \otimes \left( \Sigma^{\frac{1}{2}} \int_0^1 W(r)dW(r)'\Sigma^{\frac{1}{2}} + \Sigma_1 \right) \right) \right] \nu_p \] \tag{53}

**Proof.** Let \( \tilde{\Delta}_T \) be a sequence of \( p \times p \) positive definite matrices such that \( \tilde{\Delta}_T \xrightarrow{p} \Delta \). Consider the test statistics given by:

\[ F_{\tilde{\Delta}_T} = (\tilde{\alpha}_{\text{FPS}} - \nu_p)' \left[ \left( \Phi_{-1}' \tilde{\Sigma}^{-1} \Phi_{-1} \right) \left( \Phi_{-1}' \tilde{P}^{-1} (I_T \otimes \tilde{\Delta}_T) \tilde{P}^{-1} \Phi_{-1} \right)^{-1} \left( \Phi_{-1}' \tilde{\Sigma}^{-1} \Phi_{-1} \right) \right] \times (\tilde{\alpha}_{\text{FPS}} - \nu_p) \]

\[ F_{\Delta_T} = \left[ \nu_p \left( \sum_{t=1}^T \mu_t \tilde{\Sigma}^{-1} \phi_{t-1} \right) \left( \Phi_{-1}' \tilde{\Sigma}^{-1} \Phi_{-1} \right)^{-1} \times \left( \Phi_{-1}' \tilde{\Sigma}^{-1} \Phi_{-1} \right) \left( \Phi_{-1}' \tilde{P}^{-1} (I_T \otimes \tilde{\Delta}_T) \tilde{P}^{-1} \Phi_{-1} \right)^{-1} \left( \Phi_{-1}' \tilde{\Sigma}^{-1} \Phi_{-1} \right) \times \right. \]

\[ \left. \left( \Phi_{-1}' \tilde{\Sigma}^{-1} \Phi_{-1} \right)^{-1} \left( \sum_{t=1}^T \phi_{t-1} \tilde{\Sigma}^{-1} \mu_t \right) \nu_p \right] \]

\[ = \left[ \nu_p \left( \sum_{t=1}^T \mu_t \tilde{\Sigma}^{-1} \phi_{t-1} \right) \left( \sum_{t=1}^T \phi_{t-1} \tilde{P}^{-1} \tilde{\Delta}_T \tilde{P}^{-1} \phi_{t-1} \right)^{-1} \left( \sum_{t=1}^T \phi_{t-1} \tilde{\Sigma}^{-1} \mu_t \right) \nu_p \right] \]

Using the results in proposition 4 and taking into account that \( \tilde{P} \) is a consistent estimate of \( \Sigma^{\frac{1}{2}} \), it follows that:
\[
F_{\Delta_T} \Rightarrow \left[ \iota_p \left( \Sigma^{-1} \odot \left( \Sigma^{\frac{1}{2}} \int_0^1 dW(r)W(r)\Sigma^{\frac{1}{2}} + \Sigma'_1 \right) \right) \right] \times \\
\left( \Sigma^{-\frac{1}{2}} \Delta \Sigma^{-\frac{1}{2}} \odot \Sigma^{\frac{1}{2}} \int_0^1 W(r)W'(r)dr \Sigma^{\frac{1}{2}} \right)^{-1} \times \\
\left( \Sigma^{-1} \odot \left( \Sigma^{\frac{1}{2}} \int_0^1 W(r)dW(r)\Sigma^{\frac{1}{2}} + \Sigma'_1 \right) \right) \iota_p 
\]

The results follows from setting \( \tilde{\Delta}_T = \tilde{\Sigma}_{0RDP} \) and \( \tilde{\Delta}_T = \tilde{\Sigma}_{RDP} \). □

4 Distribution under \( H_1 \)

In the last section, the four test statistics that will be studied here were described and their asymptotic distribution under \( H_0 \) were calculated. We now turn our attention to determining their asymptotic distribution under a sequence of alternatives that converge asymptotically to the null. We will be following the theory of nearly integrated processes described for the univariate case in Phillips (1987b).

The alternative processes will be given by:

\[
z_t = A_C z_{t-1} + e_t \\
A_C = \exp \left( \frac{1}{T} C \right), \quad \text{a } p \times p \text{ matrix} \\
z_0 = 0
\]  

(54)

We will use the results presented in Phillips (1988) in order to derive the distribution of our test statistics as \( T \uparrow \infty \) when the true process is given by (54). It is important to note that here we have a sequence of alternative hypotheses which approaches \( H_0 \) as \( T \uparrow \infty \).

The following result based on Phillips (1988) is presented here since much of the discussion in this section will be based on it:

**Theorem 3** Let \( \{z_t\}_{t=1}^{\infty} \) be given by (54) and \( \{e_t\}_{t=1}^{\infty} \) be a weakly stationary sequence of \( p \)-dimensional vectors. Assume that:

1. \( E(e_t) = 0, \forall t \)

2. \( \sup_{t, \epsilon} E|e_t|^{\beta + \epsilon} < \infty \) for some \( \beta > 2 \) and \( \epsilon > 0 \)
3. \{e_t\}_{t=1}^\infty is strong mixing with mixing numbers \(\alpha_m\) that satisfy:
\[
\sum_{j=1}^\infty \alpha_m^{1-\frac{2}{n}} < \infty
\]

Define:

- \(\Sigma\) is the long run covariance matrix for process \(e_t\)
- \(J(r; C, P)\) is the Ornstein-Uhlenbeck process with matrix of coefficients \(C\) and covariance matrix \(PP'\), that is:
\[
J(r; C, P) = \int_0^r \exp(C(r-s))PdW_s, \text{ a } p \text{ dimensional vector}
\]

Under these conditions, the following results hold:

1. \(T^{-\frac{1}{2}}z_{[T/2]} \Rightarrow J(r; C, \Sigma_{\frac{1}{2}})\)
2. \(T^{-\frac{1}{2}}\sum_{t=1}^T z_t \Rightarrow \int_0^1 J(r; C, \Sigma_{\frac{1}{2}})dr\)
3. \(T^{-2}\sum_{t=1}^T z_{t-1}e_{t-1}' \Rightarrow \int_0^1 J(r; C, \Sigma_{\frac{1}{2}})J(r; C, \Sigma_{\frac{1}{2}})'dr\)
4. \(T^{-1}\sum_{t=1}^T z_{t-1}e_{t}'+e_{t}z_{t-1}' \Rightarrow \int_0^1 J(r; C, \Sigma_{\frac{1}{2}})J(1; C, \Sigma_{\frac{1}{2}}) - C \left( \int_0^1 J(r; C, \Sigma_{\frac{1}{2}})J(r; C, \Sigma_{\frac{1}{2}})'dr \right) - \left( \int_0^1 J(r; C, \Sigma_{\frac{1}{2}})J(r; C, \Sigma_{\frac{1}{2}})'dr \right) C' - \Sigma_0\)


Having enunciated this result, it is straightforward to notice that if the process \(e_t\) satisfies the same assumptions regarding the process \(u_t\) described in Assumption 1 then Theorem 3 applies to \(e_t\).

In order to derive the asymptotic distribution of the test statistics described in the previous section, it is useful to establish a result which is an analogue to Proposition 4 under \(H_1\). In order to do that, we define:

\[
\lambda_t = diag(z_t) \quad (55)
\]
\[
\nu_t = diag(e_t) \quad (56)
\]
\[
\Lambda_{-1} = \begin{bmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_{T-1} \end{bmatrix}' \quad (57)
\]
\[
\Lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_T \end{bmatrix}' \quad (58)
\]
Having defined these variables, we proceed to state a result which will be helpful in determining the asymptotic distribution of the RPD and FPS tests under the sequence of alternatives described in (54).

**Proposition 9** Suppose that \( e_t \) satisfies Assumption 1 and let \( \lambda_t \) and \( \nu_t \) be given as in (55) and (56) respectively and \( \Theta \) be a \( p \times p \) matrix. Under these conditions, we have as \( T \uparrow \infty \):

1. \( T^{-\frac{3}{2}} \left( \sum_{t=1}^{T} \lambda_{t-1} \right) \Rightarrow I_p \odot \left( \left( \int_0^1 J(r; C, \Sigma_t^{\frac{1}{2}})dr \right) \otimes \iota_p \right) \)

2. \( T^{-2} \left( \sum_{t=1}^{T} \lambda_{t-1}' \Theta \lambda_{t-1} \right) \Rightarrow \Theta \odot \left( \left( \int_0^1 J(r; C, \Sigma_t^{\frac{1}{2}})J(r; C, \Sigma_{t-1}^{\frac{1}{2}})dr \right) \right) \)

3. \( T^{-1} \left( \sum_{t=1}^{T} \nu_t' \right) = \Theta \odot \left( \int_0^1 J(r; C, \Sigma_t^{\frac{1}{2}})dW(r) \Sigma_t^{\frac{1}{2}} + \Sigma_t \right) \)

where \( \Sigma_t^{\frac{1}{2}} \) is as defined in Theorem 2.

**Proof.** From (55) and (56), it follows that:

\[
T^{-\frac{3}{2}} \sum_{t=1}^{T} \lambda_t = I_p \odot \left( T^{-\frac{3}{2}} \sum_{t=1}^{T} z_t \otimes \iota_p \right)
\]

\[
T^{-2} \sum_{t=1}^{T} \lambda_{t-1}' \Theta \lambda_{t-1} = \Theta \odot T^{-2} \sum_{t=1}^{T} z_{t-1} z_{t-1}'
\]

\[
T^{-1} \sum_{t=1}^{T} \lambda_{t-1}' \Theta \nu_t' = \Theta \odot T^{-1} \sum_{t=1}^{T} z_{t-1} u_t'
\]

Combining these and Theorem 3, the results follow from the continuous mapping theorem.

Having the last two results, the steps involved in determining the asymptotic distribution of the test statistics described in section 3 under the sequence of alternatives considered by Phillips (1987b) are very similar to those employed when determining their asymptotic behavior under \( H_0 \). In the following subsections, results concerning the distributions under any alternative of the form described in (54) are obtained. One can then determine, for each test, what is the probability that the test statistic under a certain alternative hypothesis will be in the region of acceptance of \( H_0 \), that is, the asymptotic power of the test against that particular alternative.
4.1 The PD Test under $H_1$

The following proposition provides the distributions of the test statistics of the PD procedure under the set of alternative hypotheses considered.

**Proposition 10** Let $\{z_t\}_{t=1}^{\infty}$ be given by (54) and consider the test statistics $F_{0PD}$, $F_{PD}$, $F_{0SPD}$ and $F_{SPD}$ defined in (23), (24), (27) and (28) respectively. Suppose further that $u_t$ satisfies assumption 1. Define:

$$G = C \left( \int_0^1 J(r; C, \Sigma^{\frac{1}{2}}) J(r; C, \Sigma^{\frac{1}{2}})' dr \right) + \left( \int_0^1 J(r; C, \Sigma^{\frac{1}{2}}) J(r; C, \Sigma^{\frac{1}{2}})' dr \right) C'$$

Then:

$$F_{0PD} \Rightarrow \frac{1}{4} tr \left[ \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma_0 \right) \Sigma_0^{-1} \times \right.
\left. \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma_0 \right) \times \right.$$\noindent$
\left( \int_0^1 J(r; C, \Sigma^{\frac{1}{2}}) J(r; C, \Sigma^{\frac{1}{2}})' dr \right)^{-1} \right]
$$

(59)

$$F_{PD} \Rightarrow \frac{1}{4} tr \left[ \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma_0 \right) \Sigma^{-1} \times \right.$$
\left. \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma_0 \right) \times \right.$$\noindent$
\left( \int_0^1 J(r; C, \Sigma^{\frac{1}{2}}) J(r; C, \Sigma^{\frac{1}{2}})' dr \right)^{-1} \right]
$$

(60)

$$F_{0SPD} \Rightarrow \frac{1}{4} tr \left\{ \left[ \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma \right) \Sigma^{-1} \times \right.$$
\left. \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma \right) \times \right.$$\noindent$
\left. \left( \int_0^1 J(r; C, \Sigma^{\frac{1}{2}}) J(r; C, \Sigma^{\frac{1}{2}})' dr \right)^{-1} \right\} 
$$

(61)

$$F_{SPD} \Rightarrow \frac{1}{4} tr \left\{ \left[ \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma \right) \Sigma^{-1} \times \right.$$
\left. \left( J(1; C, \Sigma^{\frac{1}{2}}) J(1; C, \Sigma^{\frac{1}{2}})' - G - \Sigma \right) \times \right.$$\noindent$
\left. \left( \int_0^1 J(r; C, \Sigma^{\frac{1}{2}}) J(r; C, \Sigma^{\frac{1}{2}})' dr \right)^{-1} \right\} 
$$

(62)
Proof. The first convergence is obtained by following the steps:

\[ F_{0PD} = tr \left[ (\hat{A}_{PD} - I) \Sigma_0^{-1} (\hat{A}_{PD} - I) \mathcal{Y}^{-1} \mathcal{Y}'^{-1} \right] \]
\[ = \frac{1}{4} tr \left[ \left( T^{-1} \sum_{t=1}^{T} y_{t-1} u'_t + u_t y'_{t-1} \right) \Sigma_0^{-1} \left( T^{-1} \sum_{t=1}^{T} y_{t-1} u'_t + u_t y'_{t-1} \right) \times \right. \]
\[ \left. \left( T^{-2} \sum_{t=1}^{T} y_{t-1} y'_{t-1} \right)^{-1} \right] \]
\[ \Rightarrow \frac{1}{4} tr \left[ \left( J(1; C, \Sigma^\frac{1}{2}) J(1; C, \Sigma^\frac{1}{2})' - G - \Sigma_0 \right) \Sigma_0^{-1} \times \right. \]
\[ \left. \left( J(1; C, \Sigma^\frac{1}{2}) J(1; C, \Sigma^\frac{1}{2})' - G - \Sigma_0 \right) \times \right. \]
\[ \left. \left( \int_0^1 J(s; C, \Sigma^\frac{1}{2}) J(s; C, \Sigma^\frac{1}{2})' ds \right)^{-1} \right] \]

We now prove the second convergence:

\[ F_{PD} = tr \left[ (\hat{A}_{PD} - I) \Sigma^{-1} (\hat{A}_{PD} - I) \mathcal{Y}^{-1} \mathcal{Y}'^{-1} \right] \]
\[ = \frac{1}{4} tr \left[ \left( T^{-1} \sum_{t=1}^{T} y_{t-1} u'_t + u_t y'_{t-1} \right) \Sigma^{-1} \left( T^{-1} \sum_{t=1}^{T} y_{t-1} u'_t + u_t y'_{t-1} \right) \times \right. \]
\[ \left. \left( T^{-2} \sum_{t=1}^{T} y_{t-1} y'_{t-1} \right)^{-1} \right] \]
\[ \Rightarrow \frac{1}{4} tr \left[ \left( J(1; C, \Sigma^\frac{1}{2}) J(1; C, \Sigma^\frac{1}{2}) - G - \Sigma_0 \right) \Sigma^{-1} \times \right. \]
\[ \left. \left( J(1; C, \Sigma^\frac{1}{2}) J(1; C, \Sigma^\frac{1}{2}) - G - \Sigma_0 \right) \times \right. \]
\[ \left. \left( \int_0^1 J(s; C, \Sigma^\frac{1}{2}) J(s; C, \Sigma^\frac{1}{2})' ds \right)^{-1} \right] \]

For the third convergence, define:

\[ C_1 = tr \left[ T^{-2} y_T y'_T (\hat{\Sigma}_{0PD}^{-1} - \hat{\Sigma}_{PD}^{-1}) y_T y'_T (T^{-2} \mathcal{Y}_{-1} \mathcal{Y}'_{-1})^{-1} \right] \]
\[ C_2 = tr \left[ (\hat{\Sigma}_{0PD} - \hat{\Sigma}_{PD}) (T^{-2} \mathcal{Y}_{-1} \mathcal{Y}'_{-1})^{-1} \right] \]

Thus

\[ C_1 \Rightarrow \frac{1}{4} tr \left[ \left( J(1; C, \Sigma^\frac{1}{2}) J(1; C, \Sigma^\frac{1}{2})' (\Sigma_0^{-1} - \Sigma^{-1}) J(1; C, \Sigma^\frac{1}{2}) J(1; C, \Sigma^\frac{1}{2})' \right) \times \right. \]
\[ \left. \left( \int_0^1 J(r; C, \Sigma^\frac{1}{2}) J(r; C, \Sigma^\frac{1}{2})' dr \right)^{-1} \right] \]
\[ C_2 \Rightarrow \frac{1}{4} tr \left[ (\Sigma_0 - \Sigma) \left( \int_0^1 J_C(r) J_C(r)' dr \right)^{-1} \right] \]
In the following, we denote $J(1; C, \Sigma^{\pm})$ by $J(1)$ and $\left( \int_0^1 J_C(r)J_C(r)dr \right)$ by $I(J)$ for simplicity. We then get that $F_{0PD} - \frac{1}{4} [C_1 + C_2]$ converges to

\[
\frac{1}{4} \text{tr} \left[ \left( J(1) J(1)' \Sigma_0^{-1} - G \Sigma_0^{-1} - I \right) (J(1) J(1)' - G - \Sigma_0) \\
- \left( J(1) J(1)' \Sigma_0^{-1} J(1) J(1)' - J(1) J(1)' \Sigma_0^{-1} J(1) J(1)' \right) - (\Sigma_0 - \Sigma) \right] (I(J))^{-1}
\]

\[
= \frac{1}{4} \text{tr} \left[ \left( J(1) J(1)' \Sigma_0^{-1} J(1) J(1)' - J(1) J(1)' \Sigma_0^{-1} J(1) J(1)' \right) \\
- G \Sigma_0^{-1} J(1) J(1)' + G \Sigma_0^{-1} G + G \\
- J(1) J(1)' + G + \Sigma_0 - J(1) J(1)' \Sigma_0^{-1} J(1) J(1)' \\
- (\Sigma_0 - \Sigma) \right] (I(J))^{-1}
\]

\[
= \frac{1}{4} \text{tr} \left[ \left( (J(1) J(1)' - G - \Sigma) \Sigma_0^{-1} (J(1) J(1)' - G - \Sigma) \right) \\
+ J(1) J(1)' (\Sigma_0^{-1} - \Sigma) J(1) J(1)' + J(1) J(1)' (\Sigma_0^{-1} - \Sigma) J(1) J(1)' \\
+ G (\Sigma_0^{-1} - \Sigma) J(1) J(1)' + G (\Sigma_0^{-1} - \Sigma) J(1) J(1)' \\
- J(1) J(1)' (\Sigma_0^{-1} - \Sigma) J(1) J(1)' \right] (I(J))^{-1}
\]

Thus, we can write

\[
F_{0PD} \Rightarrow \frac{1}{4} \text{tr} \left[ \left( (J(1; C, \Sigma^{\pm}) J(1; C, \Sigma^{\pm})' - G - \Sigma) \Sigma^{-1} \times \\
(J(1; C, \Sigma^{\pm}) J(1; C, \Sigma^{\pm})' - G - \Sigma) \\
+ (J(1; C, \Sigma^{\pm}) J(1; C, \Sigma^{\pm})' + G) (\Sigma^{-1} - \Sigma_0^{-1}) (J(1; C, \Sigma^{\pm}) J(1; C, \Sigma^{\pm})' + G) \\
- J(1; C, \Sigma^{\pm}) J(1; C, \Sigma^{\pm})' (\Sigma^{-1} - \Sigma_0^{-1}) J(1; C, \Sigma^{\pm}) J(1; C, \Sigma^{\pm})' \right) \\
\left( \int_0^1 J(r; C, \Sigma^{\pm}) J(r; C, \Sigma^{\pm})' dr \right)^{-1}
\]

Finally, we prove the fourth convergence. For that, define:

\[
C_1 = \text{tr} \left[ \left( T^{-1} y_T P_D' (\tilde{\Sigma}^{-1}_{PD} - \tilde{\Sigma}_{0PD}) \right) (T^{-2} Y_{-1} Y_{-1}^{-1})^{-1} \right]
\]

29
We then note that, by defining $K_1 = J(1)J(1)\Sigma^{-1}(\Sigma - \Sigma_0)$ and $K_2 = I(J)$ and, since $K_2$ is symmetric, $tr(K_1K_2) = tr(K_2'K_1') = tr(K_2'K_2') = tr(K_2K_2)$. Hence:

$$C_1 \Rightarrow tr \left[ \left( (\Sigma - \Sigma_0) \Sigma^{-1} J(1)J(1)' \right) (I(J))^{-1} \right]$$

We then note that $F_{SPD} = F_{PD} - \frac{3}{4}C_1 - \frac{1}{4}C_2 + \frac{2}{4}C_3$ converges to:

$$\frac{1}{4}tr \left[ \left( (J(1)J(1)'\Sigma^{-1} - G\Sigma^{-1} - \Sigma_0\Sigma^{-1} \right) (J(1)J(1)' - G - \Sigma_0) \\
- (J(1)J(1)'\Sigma^{-1}(\Sigma - \Sigma_0) - (\Sigma - \Sigma_0)\Sigma^{-1}J(1)J(1)') \\
- (\Sigma^{-1}\Sigma + \Sigma^{-1}\Sigma - \Sigma_0\Sigma^{-1}\Sigma_0) \right] (I(J))^{-1}$$

$$= \frac{1}{4}tr \left[ J(1)J(1)'\Sigma^{-1}J(1)J(1)' - J(1)J(1)'\Sigma^{-1}G - J(1)J(1)'\Sigma^{-1}\Sigma_0 \\
- G\Sigma^{-1}J(1)J(1)' + G\Sigma^{-1}G + G\Sigma^{-1}\Sigma_0 \\
- \Sigma_0\Sigma^{-1}J(1)J(1)' + \Sigma_0\Sigma^{-1}G + \Sigma_0\Sigma^{-1}\Sigma_0 \\
- J(1)J(1)'\Sigma^{-1}\Sigma + J(1)J(1)'\Sigma^{-1}\Sigma_0 - \Sigma\Sigma^{-1}J(1)J(1)' + \Sigma_0\Sigma^{-1}J(1)J(1)' \\
- \Sigma\Sigma^{-1}\Sigma + \Sigma\Sigma^{-1}\Sigma - \Sigma_0\Sigma^{-1}\Sigma_0 \right] (I(J))^{-1}$$

$$= \frac{1}{4}tr \left[ J(1)J(1)'\Sigma^{-1}J(1)J(1)' - J(1)J(1)'\Sigma^{-1}G - J(1)J(1)'\Sigma^{-1}\Sigma \\
- G\Sigma^{-1}J(1)J(1)' + G\Sigma^{-1}G + G\Sigma^{-1}\Sigma \\
- \Sigma\Sigma^{-1}J(1)J(1)' + \Sigma\Sigma^{-1}G + \Sigma\Sigma^{-1}\Sigma \\
+ G\Sigma^{-1}(\Sigma_0 - \Sigma) + (\Sigma_0 - \Sigma)\Sigma^{-1}G \right] (I(J))^{-1}$$
\[ = \frac{1}{4} \text{tr} \left[ (J(1)J(1)' - G - \Sigma) \Sigma^{-1} (J(1)J(1)' - G - \Sigma) + G \Sigma^{-1} (\Sigma_0 - \Sigma) \\
+ (\Sigma_0 - \Sigma) \Sigma^{-1} G \right] (I(J))^{-1} \]

Therefore, we can write:

\[ F_{SPD} \Rightarrow \frac{1}{4} \text{tr} \left\{ \left[ (J(1; C, \Sigma^{-\frac{1}{2}})) J(1; C, \Sigma^{-\frac{1}{2}})' - G - \Sigma \right] \Sigma^{-1} \times \\
\left( J(1; C, \Sigma^{-\frac{1}{2}})) J(1; C, \Sigma^{-\frac{1}{2}})' - G - \Sigma \right) \\
+ G \Sigma^{-1} (\Sigma_0 - \Sigma) + (\Sigma_0 - \Sigma) \Sigma^{-1} G \right\} \times \\
\left( \int_0^1 J(r; C, \Sigma^{-\frac{1}{2}}) J(r; C, \Sigma^{-\frac{1}{2}})' dr \right)^{-1} \]

One important fact given these asymptotic distributions is that, under the set of alternative hypotheses considered, the \( F_{0SPD} \) and \( F_{SPD} \) are not free of nuisance parameters contrary to what happens under the null. This suggests that the power of the test may be affected by the nuisance parameters, despite the corrections made in order to free its distributions under \( H_0 \) from them.

\subsection*{4.2 The RPD Test under \( H_1 \)}

The following proposition provides the distributions of the test statistics of the RPD procedure under the set of alternative hypotheses considered.

\textbf{Proposition 11} Let \( \{z_t\}_{t=1}^\infty \) be given by (54) and consider the test statistics \( F_{0RPD} \) and \( F_{RPD} \) defined in (42) and (43) respectively. Suppose further that \( u_t \) satisfies assumption 1. Then:

\[ F_{0RPD} \Rightarrow \left[ t_p \left( I_p \odot (\Sigma_1^{\frac{1}{2}}) \right) \int_0^1 dW(r) J(r; C; \Sigma_1^{\frac{1}{2}}) J(r; C; \Sigma_1^{\frac{1}{2}})' \right] \times \\
\left( \int_0^1 J(r; C; \Sigma_1^{\frac{1}{2}}) J(r; C; \Sigma_1^{\frac{1}{2}})' dr \right)^{-1} \times \\
\left( I_p \odot (\int_0^1 J(r; C; \Sigma_1^{\frac{1}{2}}) dW(r) \Sigma_1^{\frac{1}{2}} + \Sigma_1) \right) t_p \] \hfill (63)
\[ F_{RPD} \Rightarrow \left[ t_p' \left( I_p \otimes \left( \Sigma^\frac{1}{2} \int_0^1 dW(r)J(r; C, \Sigma^\frac{1}{2})' + \Sigma_i' \right) \right) \times \left( \Sigma \otimes \int_0^1 J(r; C, \Sigma^\frac{1}{2})J(r; C, \Sigma^\frac{1}{2})' \right)^{-1} \times \left( I_p \otimes \left( \int_0^1 J(r; C, \Sigma^\frac{1}{2})dW(r)\Sigma^\frac{1}{2} + \Sigma_i \right) \right) t_p \right] \]  

(64)

\textbf{Proof.} Let \( \tilde{\Delta}_T \) be a sequence of \( p \times p \) positive definite matrices such that \( \tilde{\Delta}_T \rightarrow_{P} \Delta \). Consider the test statistics given by:

\[ F_{\Delta_T} = (\tilde{\alpha}_{RPD} - \iota_p) \left[ (\Phi'_{-1} \Phi_{-1}) \left[ \Phi'_{-1} \left( I_T \otimes \tilde{\Delta}_T \right) \Phi_{-1} \right]^{-1} \left( \Phi'_{-1} \Phi_{-1} \right) \right] (\tilde{\alpha}_{RPD} - \iota_p) \]

\[ = \left[ t_p' \left( \sum_{t=1}^T \mu_t \phi_{t-1}' \right) \left( \Phi'_{-1} \Phi_{-1} \right)^{-1} \times \left( \Phi'_{-1} \Phi_{-1} \right) \left( \Phi'_{-1} \left( I_T \otimes \tilde{\Delta}_T \right) \Phi_{-1} \right)^{-1} \left( \Phi'_{-1} \Phi_{-1} \right) \times \left( \Phi'_{-1} \Phi_{-1} \right)^{-1} \left( \sum_{t=1}^T \phi_{t-1} \mu_t \right) t_p \right] \]

\[ = \left[ t_p' \left( I_p \otimes \sum_{t=1}^T u_t \phi_{t-1}' \right) \left( \tilde{\Delta}_T \otimes \sum_{t=1}^T y_{t-1} \phi_{t-1}' \right)^{-1} \left( I_p \otimes \sum_{t=1}^T y_{t-1} \mu_t \right) t_p \right] \]

Using Proposition 9 it follows that:

\[ F_{\Delta_T} \Rightarrow \left[ t_p' \left( I_p \otimes \left( \Sigma^\frac{1}{2} \int_0^1 dW(r)J(r; C, \Sigma^\frac{1}{2})' + \Sigma_i' \right) \right) \times \left( \Delta \otimes \int_0^1 J(r; C, \Sigma^\frac{1}{2})J(r; C, \Sigma^\frac{1}{2})'dr \right)^{-1} \times \left( I_p \otimes \left( \int_0^1 J(r; C, \Sigma^\frac{1}{2})dW(r)\Sigma^\frac{1}{2} + \Sigma_i \right) \right) t_p \right] \]

The results follow from setting \( \tilde{\Delta}_T = \tilde{\Sigma}_{0RPD} \) and \( \tilde{\Delta}_T = \tilde{\Sigma}_{RPD} \).

\[ \text{4.3 The FPS Test under } H_1 \]

The following proposition provides the distributions of the test statistics of the FPS procedure under the set of alternative hypotheses considered.

32
Proposition 12 Let \( \{z_i\}_{i=1}^\infty \) be given by (54) and consider the test statistics \( F_{0FPS} \) and \( F_{FPS} \) defined in (50) and (51) respectively. Suppose further that \( u_t \) satisfies assumption 1. Then:

\[
F_{0FPS} \Rightarrow \left[ t'_p \left( \Sigma^{-1} \odot (\Sigma_{\frac{3}{2}} \int_0^1 dW(r)J(r; C, \Sigma_{\frac{3}{2}}') + \Sigma_1') \right) \right. \\
\left. \left( \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}} \odot \int_0^1 J(r; C, \Sigma_{\frac{3}{2}})J(r; C, \Sigma_{\frac{3}{2}}') \right)^{-1} \times \\
\left( \Sigma^{-1} \odot (\int_0^1 J(r; C, \Sigma_{\frac{3}{2}}) dW(r)') \Sigma_{\frac{3}{2}} + \Sigma_1) \right) t_p \right]
\]

\[
F_{FPS} \Rightarrow \left[ t'_p \left( \Sigma^{-1} \odot (\Sigma_{\frac{3}{2}} \int_0^1 dW(r)J(r; C, \Sigma_{\frac{3}{2}}') + \Sigma_1') \right) \times \\
\left( I_p \odot \int_0^1 J(r; C, \Sigma_{\frac{3}{2}})J(r; C, \Sigma_{\frac{3}{2}}') \right)^{-1} \times \\
\left( \Sigma^{-1} \odot (\int_0^1 J(r; C, \Sigma_{\frac{3}{2}}) dW(r)') \Sigma_{\frac{3}{2}} + \Sigma_1) \right) t_p \right]
\]

Proof. Let \( \tilde{\Delta}_T \) be a sequence of \( p \times p \) positive definite matrices such that \( \tilde{\Delta}_T \xrightarrow{p} \Delta \) and \( \tilde{P} \) be a positive definite consistent estimator of \( \Sigma_{\frac{3}{2}} \). Consider the test statistics given by:

\[
F_{\Delta_T} = (\tilde{\alpha}_{FPS} - t_p)' \left[ (\Phi'_{-1} \Phi_{-1}) \left[ \Phi'_{-1} \tilde{P}^{-1'} (I_T \odot \tilde{\Delta}_T) \tilde{P}^{-1} \Phi_{-1} \right]^{-1} (\Phi'_{-1} \Phi_{-1}) \right] (\tilde{\alpha}_{FPS} - t_p)
\]

\[
= \left[ t'_p \left( \sum_{t=1}^T \mu_t' \tilde{\Sigma}^{-1} \phi_{t-1} \right) \left( \Phi'_{-1} \tilde{\Sigma}^{-1} \Phi_{-1} \right)^{-1} \times \\
\left( \Phi'_{-1} \tilde{\Sigma}^{-1} \Phi_{-1} \right) \left( \Phi'_{-1} \tilde{P}^{-1'} (I_T \odot \tilde{\Delta}_T) \tilde{P}^{-1} \Phi_{-1} \right)^{-1} \left( \Phi'_{-1} \tilde{\Sigma}^{-1} \Phi_{-1} \right) \times \\
\left( \Phi'_{-1} \tilde{\Sigma}^{-1} \Phi_{-1} \right)^{-1} \left( \sum_{t=1}^T \phi_{t-1}' \tilde{\Sigma}^{-1} \mu_t \right) t_p \right] \\
= \left[ t'_p \left( \sum_{t=1}^T \mu_t' \tilde{\Sigma}^{-1} \phi_{t-1} \right) \left( \sum_{t=1}^T \phi_{t-1}' \tilde{P}^{-1'} \tilde{\Delta}_T \tilde{P}^{-1} \phi_{t-1} \right)^{-1} \left( \sum_{t=1}^T \phi_{t-1}' \tilde{\Sigma}^{-1} \mu_t \right) t_p \right]
\]

Using Proposition 9 it follows that:

\[
F_{\Delta_T} \Rightarrow \left[ t'_p \left( \Sigma^{-1} \odot (\Sigma_{\frac{3}{2}} \int_0^1 dW(r)J(r; C, \Sigma_{\frac{3}{2}}') + \Sigma_1') \right) \times \\
\left( \Sigma^{-\frac{1}{2}} \Delta \Sigma^{-\frac{1}{2}} \odot \int_0^1 J(r; C, \Sigma_{\frac{3}{2}})J(r; C, \Sigma_{\frac{3}{2}}') dr \right)^{-1} \times \\
\left( \Sigma^{-1} \odot (\int_0^1 J(r; C, \Sigma_{\frac{3}{2}}) dW(r)') \Sigma_{\frac{3}{2}} + \Sigma_1) \right) t_p \right]
\]

The results follow from setting \( \tilde{\Delta}_T = \tilde{\Sigma}_{0RDP} \) and \( \tilde{\Delta}_T = \tilde{\Sigma}_{RDP} \).
5 Conclusion

In this essay, the distributions under a certain set of alternative hypotheses of the multivariate unit roots tests proposed in Phillips & Durlauf (1986) and Flóres, Preumont & Szafarz (1996) were determined. In addition to that, a new test is proposed in which the restriction employed in the Flóres, Preumont & Szafarz test is combined with the OLS estimation used in the Phillips & Durlauf test and its distributions both under the null of unit roots and the same set of alternative hypotheses are determined.

The family of alternative hypotheses considered may be regarded as the multivariate version of the set of alternatives considered in Phillips (1987b). One interesting result emerging from the analysis made is the fact that the corrections made in order to free the asymptotic distribution of the Phillips & Durlauf test under the null of nuisance parameters are not able to do the same under the set of alternative hypotheses considered. This strongly suggests that the power of the test against alternatives in this set is affected by some nuisance parameters, including the long run covariance matrix.

Knowing the asymptotic distribution under the null hypothesis and a set of alternatives allows us to determine the asymptotic power of the tests in order to compare them. With the new test proposed, one may be able not only to compare the power of each test but to identify what causes the difference in performance.

The tests dealt with in this essay involve multiple parameters and a Wald test was used in the three cases in order to reach a test statistic. Since the asymptotic distributions of the tests considered are not normal, this may lead to some loss of power. One alternative to that is to determine a minimum area multivariate region of acceptance for a given test size. This, however, may prove to be a gruesome task. One promising way to tackle this problem consists in trying to adapt the results in Bucchianico, Einmahl & Mushkudiani (2001) to the problem at hand.
References


A Estimation of unrestricted VAR(1) coefficients

Consider the model given by:

\[ y_t = Ay_{t-1} + u_t \] (A- 1)

where \( y_t \) is a \( p \)-dimensional vector. Define the following arrays:

\[
\mathcal{Y}_{-1} = \begin{bmatrix} y_0 & y_1 & \cdots & y_{T-1} \end{bmatrix} \text{ a } p \times T \text{ matrix}
\]

\[
Y_{-1} = vec(\mathcal{Y}_{-1}) \text{ a } Tp\text{-dimensional vector}
\]

\[
\mathcal{Y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_T \end{bmatrix} \text{ a } p \times T \text{ matrix}
\]

\[
Y = vec(\mathcal{Y}) \text{ a } Tp\text{-dimensional vector}
\]

Since \( Ay_{t-1} \) is a column vector, it follows that \( Ay_{t-1} = vec(Ay_{t-1}) = vec(y'_{t-1}A') \). Therefore, we may estimate the coefficient matrix \( A \) using OLS in two different versions of the same equation:

\[
vec(y_t) = vec(Ay_{t-1}) + vec(u_t)
\]

= \left[ y'_{t-1} \otimes I_p \right] vec(A) + vec(u_t) \quad (A- 2)

\[
vec(y_t) = vec(y'_{t-1}A') + vec(u_t)
\]

= \left[ I_p \otimes y'_{t-1} \right] vec(A') + vec(u_t) \quad (A- 3)

Stacking (A- 2) and (A- 3) from \( t = 1 \) to \( t = T \), we can write:

\[
Y = [\mathcal{Y}_{-1} \otimes I_p](vec(A) + U_t) \quad (A- 4)
\]

Applying OLS to (A- 4), we get:

\[
vec(\hat{A}) = \left\{ \left[ \mathcal{Y}_{-1} \otimes I_p \right]' \left[ \mathcal{Y}_{-1} \otimes I_p \right] \right\}^{-1} \left\{ \left[ \mathcal{Y}_{-1} \otimes I_p \right]' Y \right\}
\]

= \left[ \left( \mathcal{Y}_{-1} \mathcal{Y}'_{-1} \right)^{-1} \otimes I_p \right] \left\{ \left[ \mathcal{Y}'_{-1} \otimes I_p \right]' vec(\mathcal{Y}) \right\}
\]

= \left[ \left( \mathcal{Y}_{-1} \mathcal{Y}'_{-1} \right)^{-1} \otimes I_p \right] vec(\mathcal{Y} \mathcal{Y}'_{-1})
\]

= vec \left( (\mathcal{Y} \mathcal{Y}'_{-1})(\mathcal{Y}_{-1} \mathcal{Y}'_{-1})^{-1} \right)
\]

\[
\hat{A} = (\mathcal{Y} \mathcal{Y}'_{-1})(\mathcal{Y}_{-1} \mathcal{Y}'_{-1})^{-1} \quad (A- 5)
\]

37
B A Lengthy Algebra Result

The following result requires some lengthy calculations to be reached and will be useful to prove some propositions:

Proposition A-I

\[
(W(1)W(1)' - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}}) \left( \Sigma^\frac{1}{2}\Sigma_0^{-1}\Sigma^\frac{1}{2} \right) \left( W(1)W(1)' - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}} \right) = \\
\left[ (W(1)W(1)' - I_p)(W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma_0 - \Sigma))\Sigma^{-\frac{1}{2}} \right] + \\
W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}W(1)W(1)'
\]

Proof.

\[
(W(1)W(1)' - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}}) \left( \Sigma^\frac{1}{2}\Sigma_0^{-1}\Sigma^\frac{1}{2} \right) \left( W(1)W(1)' - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}} \right) = \\
(W(1)W(1)')\Sigma^\frac{1}{2}\Sigma_0^{-1}\Sigma^\frac{1}{2} - I_p) \left( W(1)W(1)' - I_p + I_p - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}} \right) = \\
(W(1)W(1)'(I_p - I_p + \Sigma^\frac{1}{2}\Sigma_0^{-1}\Sigma^\frac{1}{2}) - I_p) \left( (W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0))\Sigma^{-\frac{1}{2}} \right) = \\
(W(1)W(1)'(I_p + \Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}) - I_p) \left( (W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0))\Sigma^{-\frac{1}{2}} \right) = \\
\left( (W(1)W(1)' - I_p) + W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2} \right) \times \\
\left( (W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0))\Sigma^{-\frac{1}{2}} \right) = \\
\left[ (W(1)W(1)' - I_p)(W(1)W(1)' - I_p) + (W(1)W(1)' - I_p)(\Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0))\Sigma^{-\frac{1}{2}} \right] + \\
W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}W(1)W(1)' - W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2} + \\
+ \ W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2} - \Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0)\Sigma^{-\frac{1}{2}} \right] = \\
\left[ (W(1)W(1)' - I_p)(W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma_0 - \Sigma))\Sigma^{-\frac{1}{2}} \right] + \\
W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}W(1)W(1)' + \\
+ \ W(1)W(1)' \times \\
\left( \Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0)\Sigma^{-\frac{1}{2}} - \Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2} \right) \Sigma^{-\frac{1}{2}}(\Sigma - \Sigma_0)\Sigma^{-\frac{1}{2}} \right] = \\
\left[ (W(1)W(1)' - I_p)(W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma_0 - \Sigma))\Sigma^{-\frac{1}{2}} \right] + \\
W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}W(1)W(1)' + \\
W(1)W(1)' \left( (I_p - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}}) + (\Sigma^\frac{1}{2}\Sigma_0^{-1}\Sigma^\frac{1}{2} - I_p)(-I_p + I_p - \Sigma^{-\frac{1}{2}}\Sigma_0\Sigma^{-\frac{1}{2}}) \right) = \\
\left[ (W(1)W(1)' - I_p)(W(1)W(1)' - I_p) + (\Sigma^{-\frac{1}{2}}(\Sigma_0 - \Sigma))\Sigma^{-\frac{1}{2}} \right] + \\
W(1)W(1)'\Sigma^\frac{1}{2}(\Sigma_0^{-1} - \Sigma^{-1})\Sigma^\frac{1}{2}W(1)W(1)'
\]
Consider the process given by:

\[ y_t = \alpha y_{t-1} + u_t \]
\[ \alpha = \exp \left[ T^{-1}C \right] \]

where \( y_t \) and \( u_t \) are \( p \)-dimensional vectors and \( C \) is a \( p \times p \) matrix.

Define:

\[ S_T(r) = \sum_{i=1}^{[Tr]} u_i \]
\[ X_T(r) = T^{-\frac{1}{2}} PS_{[Tr]} = T^{-\frac{1}{2}} P S_{j-1}, \quad \frac{(j-1)}{T} \leq r < \frac{j}{T} \]

where the \( p \times p \) matrix \( P \) is the inverse of the square root of the long run covariance matrix of the process \( u_t \).

One may write:

\[ X_T\left( \frac{j}{T} \right) = T^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} S_j \]
\[ X_T\left( \frac{j-1}{T} \right) = T^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} S_{j-1} \]
\[ \int_{\frac{j-1}{T}}^{\frac{j}{T}} dX_T(s) = X_T\left( \frac{j}{T} \right) - X_T\left( \frac{j-1}{T} \right) \]
\[ = T^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} (S_j - S_{j-1}) \]
\[ = T^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} u_j \]

Writing:

\[ y_{[Tr]} = \sum_{i=1}^{[Tr]} \alpha^{(t-i)} u_i + \alpha^t y_0 \]
\[ = \sum_{i=1}^{[Tr]} \exp \left( \frac{t-i}{T} C \right) u_i + \alpha^t y_0 \]
\begin{align*}
T^{-\frac{1}{2}} y^{[Tr]} &= \sum_{i=1}^{[Tr]} \int_{t-iT}^{t-iT} \exp \left( \frac{t-iT}{C} \right) \int_{-1}^{1} dX_T(s) + \alpha^t y_0 \\
T^{-\frac{1}{2}} y^{[Tr]} &= \sum_{i=1}^{[Tr]} \int_{t-iT}^{t-iT} \exp \left( [(r-s)C] \sum^{-\frac{1}{2}} dX_T(s) + T^{-\frac{1}{2}} \alpha^t y_0 \\
&= \int_{0}^{r} \exp \left( [(r-s)C] \sum^{-\frac{1}{2}} dX_T(s) + O \left( T^{-\frac{1}{2}} \right) \\
\Rightarrow \quad P \int_{0}^{r} \exp \left( [(r-s)C] \sum^{-\frac{1}{2}} dW(s) \\
\Rightarrow \quad J(r; C, \sum^{-\frac{1}{2}}) \\
\text{Using the continuous mapping theorem:}
\end{align*}

\begin{align*}
T^{-1} y^{[Tr]} y'^{[Tr]} &\Rightarrow J(r, C, \sum^{\frac{1}{2}})J(r, C, \sum^{\frac{1}{2}})' \\
T^{-\frac{2}{2}} \sum_{i=1}^{[Tr]} y_i &\Rightarrow \int_{0}^{r} J(s, C, \sum^{\frac{1}{2}})ds \\
T^{-2} \sum_{i=1}^{[Tr]} y_i y'_i &\Rightarrow \int_{0}^{r} J(s, C, \sum^{\frac{1}{2}})J(s, C, \sum^{\frac{1}{2}})'ds
\end{align*}