Cointegration in a Periodic Vector Autoregression

Frank Kleibergen* Philip Hans Franses†

Econometric Institute Report 9906/A

Abstract

We consider representation, estimation and inference on cointegration in a periodic vector autoregressive time series model (PVAR). We show that cointegration amounts to a restriction on a product of parameter matrices. We therefore use GMM to construct estimators of the long-run (cointegration) parameters and to obtain test statistics for cointegration. We show that the limiting distributions of the GMM estimators and the corresponding test statistics in a PVAR are identical to those of the maximum likelihood cointegration estimators and test statistics in standard nonperiodic VAR models.

1. Introduction

Two key features of many seasonally observed macroeconomic time series are their trending behavior and marked seasonal fluctuations. Ample empirical evidence suggests that the trend in such data is often stochastic. Also, several empirical studies find that seasonal fluctuations evolve over time and hence that perhaps descriptive models with stochastic seasonal components are useful. For univariate time series, the two classes of models frequently considered for describing data with both evolving trends and seasonality are the seasonal integration model, see Hylleberg *et. al.* (1990) amongst others, and the periodic integration model see Franses (1996) and Boswijk, Franses and Haldrup (1997) amongst others. The univariate seasonal integration model has been extended to multivariate cointegrated time series in Lee (1992) and Johansen and Schaumburg (1998). It is the purpose of the present paper to consider a multivariate extension of the second class of models, that is, we consider cointegration in a periodic vector autoregression.

The outline of the paper is as follows. In section 2, we define cointegration in a periodic vector autoregressive model of order 1 and derive the implied stochastic trends. In section 3, we specify its log-likelihood such that it reflects the cointegration parameters. It appears impossible to obtain analytical expressions of the maximum likelihood estimators of the cointegration parameters. We therefore use the log-likelihood to construct a (quasi-log likelihood) objective function that we optimize using the generalized method of moments (GMM), see

*Econometric Institute, Erasmus University Rotterdam. P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, Email:kleibergen@few.eur.nl. The author thanks the Netherlands Organisation for Scientific Research (N.W.O.) for their financial support.

†Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, Email:franses@few.eur.nl (Corresponding author).

Hansen (1982). In section 4, we construct the limiting distributions of the cointegrating vector estimator and of the objective function. Finally, section 5 shows how our results for the first order periodic vector autoregressive model extend to higher order periodic vector autoregressions and to models with additional deterministic components. We show that only a specific representation of the higher order dynamics leads to tractable long-run properties.

2. The Periodic VAR(1)

In this section we consider a periodic VAR(1) model, which we will extend in section 5. We discuss issues concerning representation and inference on cointegration and stochastic trends.

2.1. Representation

Consider a multivariate time series y_n of dimension k, $n = 1, \ldots, N$, with seasonal frequency S, that is, there are S observations on y_n every year, where we assume that S divides N. A periodic vector autoregressive model of order 1 (PVAR(1)) for this series can be represented by,

$$y_n = \varphi_s y_{n-1} + u_n, \qquad s = 1, \ldots, S;\ n = 1, \ldots, N, \tag{2.1}$$

see Franses (1996) and Lütkepohl (1991), where φ_s is a $(k \times k)$ parameter matrix of full rank with values that can vary across the seasons, and u_n is a vector of dimension k with disturbances. The disturbances u_n are assumed to be i.i.d. with mean zero and covariance matrix Ω_s, which can thus take values that differ across the seasons. Because of these periodic parameters, the PVAR(1) likely implies different short-run properties of y_n across the seasons. The long-run properties implied by the PVAR(1) model are however identical for all seasons, as we will demonstrate later in this section.

When we recursively substitute the lagged variable in the PVAR(1) model, we can determine the relationship between a current observation y_n and the observation in the same season in the previous year, y_{n-S}. As both observations concern the same season, they obviously have identical seasonal components. The link between y_n and y_{n-S}, given (2.1), is

$$y_n = \left(\prod_{i=1}^{s} \varphi_{s-i+1} \prod_{i=1}^{S-s} \varphi_{S-i+1} \right) y_{n-S} + u_n + \sum_{j=1}^{S-1} \left(\prod_{i=1}^{j} \varphi_{s-i+1} \prod_{i=1, j>s}^{S-j-1} \varphi_{S-i+1} \right) u_{n-j}, \tag{2.2}$$

where $\prod_{i=1}^{s} \varphi_{s-i+1} \prod_{i=1}^{S-s} \varphi_{S-i+1} = (\varphi_s \cdots \varphi_1) (\varphi_S \cdots \varphi_{s+1})$. For later purposes it is convenient to write (2.2) in its nonseasonal annual format, see Tiao and Grupe (1980) and Osborn (1991). This results in

$$\Delta Y_{s,t} = \Pi_s Y_{s,t-1} + \varepsilon_{s,t}, \qquad s = 1, \ldots, S;\ t = 1, \ldots, T = \frac{N}{S}; \tag{2.3}$$

where the indices s and t in $Y_{s,t}$ indicate the season and year of an observation of y_n, that is $Y_{s,t} = y_n$, with $n = St + s$. Furthermore, (2.3) contains $\Delta Y_{s,t} = Y_{s,t} - Y_{s,t-1}$ and defines

$$\varepsilon_{s,t} = u_n + \sum_{j=1}^{S-1} \left(\prod_{i=1}^{j} \varphi_{s-i+1} \prod_{i=1, j>s}^{S-j-1} \varphi_{S-i+1} \right) u_{n-j}, \tag{2.4}$$

$n = St + s$; and

$$\Pi_s = \left(\prod_{i=1}^{s} \varphi_{s-i+1} \prod_{i=1}^{S-s} \varphi_{S-i+1} \right) - I_k. \tag{2.5}$$

.

2.2. Cointegration and Stochastic Trends

In case there is cointegration amongst the elements of y_n, the (seasonally varying) long-run relation can be obtained from

$$\Pi_s = \alpha_s \beta_s', \qquad s = 1, \ldots, S; \tag{2.6}$$

see *e.g.* Engle and Granger (1987), where α_s, β_s are $k \times r$ matrices. To identify the elements of α_s and β_s, we assume that $\beta_s = (I_r \; -\beta_{2s}')'$, with β_{2s} a $(k-r) \times r$ matrix.

When α_s and β_s do not span the same vector spaces, that is $\alpha_{s,\perp}' \beta_{s,\perp}$ has full rank for all s, where $\beta_{s,\perp}' \beta_s \equiv 0$, $\alpha_{s,\perp}' \alpha_s \equiv 0$, (2.3) can be written in the stochastic trend specification for each season, that is

$$Y_{s,t} = \beta_{s,\perp} (\alpha_{s,\perp}' \beta_{s,\perp})^{-1} \alpha_{s,\perp}' \sum_{i=1}^{t} \varepsilon_{s,i} + X_{s,t}, \qquad s = 1, \ldots, S, \tag{2.7}$$

see Johansen (1991), where $X_{s,t}$ is a stationary time series. This stochastic trend specification directly shows that $\beta_s' Y_{s,t}$ is stationary, and hence that β_s contains the cointegrating vectors with the stationary linear combinations of the k elements of y_n.

An interesting property of the PVAR(1) model is that it imposes the same long-run properties for each season. To show this, we consider the stochastic trend for season S, the last season of the year. The definition of $\varepsilon_{s,t}$ in (2.4) implies that

$$\begin{aligned} \sum_{i=1}^{t} \varepsilon_{S,i} &= \varphi_S \sum_{i=1}^{t} \varepsilon_{S-1,i} + u_{St+S}, \\ &= \varphi_S \varphi_{S-1} \sum_{i=1}^{t} \varepsilon_{S-2,i} + u_{St+S} + \varphi_S u_{St+S-1} \\ &= \left(\prod_{i=1}^{S} \varphi_{S-i+1} \right) \sum_{i=1}^{t} \varepsilon_{1,i} + u_{St+S} + \sum_{j=1}^{S-1} \left(\prod_{i=S-j}^{S} \varphi_i \right) u_{St+S-j}. \end{aligned} \tag{2.8}$$

As φ_s are of full rank and given (2.5), it holds that

$$\begin{aligned} \Pi_s &= \alpha_s \beta_s' \Leftrightarrow \\ \varphi_{s+1} \Pi_s \varphi_{s+1}^{-1} &= \varphi_{s+1} \left(\varphi_s \cdots \varphi_1 \right) \left(\varphi_S \cdots \varphi_{s+1} \right) \varphi_{s+1}^{-1} - \varphi_{s+1} \varphi_{s+1}^{-1} \Leftrightarrow \\ \varphi_{s+1} \Pi_s \varphi_{s+1}^{-1} &= \left(\varphi_{s+1} \cdots \varphi_1 \right) \left(\varphi_S \cdots \varphi_s \right) - I_k \Leftrightarrow \\ \varphi_{s+1} \Pi_s \varphi_{s+1}^{-1} &= \Pi_{s+1} \Leftrightarrow \\ \varphi_{s+1} \alpha_s \beta_s' \varphi_{s+1}^{-1} &= \alpha_{s+1} \beta_{s+1}'. \end{aligned} \tag{2.9}$$

This result implies that the spaces spanned by α_{s+1} and β_{s+1} are identical to the spaces spanned by $\varphi_{s+1} \alpha_s$ and $\varphi_{s+1}^{-1\prime} \beta_s$. As a consequence, $\alpha_{s+1,\perp} = \left(\varphi_{s+1} \alpha_s \right)_\perp = \varphi_{s+1}^{-1\prime} \alpha_{s,\perp}$ and $\beta_{s+1,\perp} =$

$\left(\varphi_{s+1}^{-1\prime}\beta_s\right)_\perp = \varphi_{s+1}\beta_{s,\perp}$. Substituting this in the stochastic trend of $Y_{s+1,t}$, and using the results in (2.8)-(2.9), results in

$$\begin{aligned}
& \beta_{s+1,\perp}(\alpha'_{s+1,\perp}\beta_{s+1,\perp})^{-1}\alpha'_{s+1,\perp}\sum_{i=1}^{t}\varepsilon_{s+1,i} \qquad (2.10)\\
= \;& \beta_{s+1,\perp}(\alpha'_{s+1,\perp}\beta_{s+1,\perp})^{-1}\alpha'_{s+1,\perp}\left(\varphi_{s+1}\sum_{i=1}^{t}\varepsilon_{s,i}+u_{St+S}\right)\\
= \;& \varphi_{s+1}\beta_{s,\perp}(\alpha'_{s,\perp}\varphi_{s+1}^{-1}\varphi_{s+1}\beta_{s,\perp})^{-1}\alpha'_{s,\perp}\varphi_{s+1}^{-1}\left(\varphi_{s+1}\sum_{i=1}^{t}\varepsilon_{s,i}+u_{St+S}\right)\\
= \;& \varphi_{s+1}\beta_{s,\perp}(\alpha'_{s,\perp}\beta_{s,\perp})^{-1}\alpha'_{s,\perp}\sum_{i=1}^{t}\varepsilon_{s,i}+\varphi_{s+1}\beta_{s,\perp}(\alpha'_{s,\perp}\beta_{s,\perp})^{-1}\alpha'_{s,\perp}\varphi_{s+1}^{-1}u_{St+S}.
\end{aligned}$$

This expression shows that the stochastic trend of $Y_{s+1,t}$ equals φ_{s+1} times the stochastic trend of $Y_{s+1,t}$ plus a stationary term that vanishes asymptotically when we divide by the square root of the number of years T. The limiting behavior of $\frac{1}{\sqrt{T}}Y_{s,t}$, $s=1,\ldots,S$, therefore concerns the same Brownian motion, for each $s=1,\ldots,S$, that is,

$$\begin{aligned}
\frac{1}{\sqrt{T}}Y_{s,t} \;&\Rightarrow\; \frac{1}{\sqrt{T}}\varphi_s Y_{s-1,t} \qquad\qquad s=1,\ldots,S; \qquad (2.11)\\
&\Rightarrow\; \frac{1}{\sqrt{T}}\left(\prod_{i=1}^{s-1}\varphi_{s-i+1}\right)Y_{1,t},\\
&\Rightarrow\; \left(\prod_{i=1}^{s-1}\varphi_{s-i+1}\right)\Phi_1^{\frac{1}{2}}B_{k-r}\\
&\Rightarrow\; (\varphi_s\cdots\varphi_2)\beta_{1,\perp}\Phi_1^{\frac{1}{2}}B_{k-r},
\end{aligned}$$

where "$\Rightarrow$" denotes weak convergence, see Billingsley (1986), where B_{k-r} is a $k-r$ dimensional Brownian motion with identity covariance matrix, and where

$$\begin{aligned}
\Phi_1 \;=\; & (\alpha'_{1,\perp}\beta_{1,\perp})^{-1}\alpha'_{1,\perp} \qquad (2.12)\\
& \left\{\Omega_1+\varphi_1\left[\sum_{j=1}^{S-1}\left(\prod_{i=1}^{j-1}\varphi_{S-i+1}\right)\Omega_{S-j+1}\left(\prod_{i=1}^{j-1}\varphi_{S-i+1}\right)'\right]\varphi_1'\right\}\\
& \alpha_{1,\perp}(\alpha'_{1,\perp}\beta_{1,\perp})^{-1\prime},
\end{aligned}$$

see Phillips and Durlauf (1986).

3. Inference on the Cointegration Rank

As can be observed from (2.9), the cointegration parameters in Π_s imply nonlinear restrictions on the φ_s parameters. Estimation and inference on these cointegration parameters is therefore not straightforward. In this section, we construct methods for estimation of and inference on the cointegration parameters in a PVAR(1). We will use an objective function, that is based on the log of the likelihood function, which we will optimize to obtain estimators of the cointegration parameters. We will show that this objective function can also be used to test for the number of cointegrating relations.

3.1. The Objective Function

To construct the log-likelihood, it is notationally convenient to specify the PVAR(1) model in (2.1) as

$$Y_{s,t} = \varphi_s Z_{s,t} + \xi_{s,t}, \qquad s = 1, \ldots, S;\ t = 1, \ldots, T, \tag{3.1}$$

where $Z_{s,t} = Y_{s-1,t}$ for $s = 2, \ldots, S$, and $Z_{1,t} = Y_{S,t-1}$. Obviously, $\xi_{s,t} = u_n$, with $n = St + s$. When we assume that the disturbances $\xi_{s,t}$ are independently normally distributed with mean zero and covariance matrix Ω_s, that is $\xi_{s,t} \sim n(0, \Omega_s)$, we can construct the following (log-) likelihood;

$$\begin{aligned} L(\varphi_1, \ldots, \varphi_S, \Omega_1, \ldots, \Omega_S) &= -\frac{1}{2}T\sum_{s=1}^{S}\log|\Omega_s| - \frac{1}{2}\sum_{s=1}^{S}\sum_{t=1}^{T}\xi'_{s,t}\Omega_s^{-1}\xi_{s,t} \\ &= -\frac{1}{2}T\sum_{s=1}^{S}\log|\Omega_s| - \frac{1}{2}\sum_{s=1}^{S} tr\left(\sum_{t=1}^{T}\Omega_s^{-1}\xi_{s,t}\xi'_{s,t}\right). \end{aligned} \tag{3.2}$$

From (3.1), we can estimate φ_s, $s = 1, \ldots, S$, by least squares, to obtain

$$\hat{\varphi}_s = \left(\sum_{t=1}^{T} Y_{s,t}Z'_{s,t}\right)\left(\sum_{t=1}^{T} Z_{s,t}Z'_{s,t}\right)^{-1},\ s = 1, \ldots, S, \tag{3.3}$$

and we can use this estimator to obtain an estimator of Ω_s, given by

$$\hat{\Omega}_s = \frac{1}{T}\sum_{t=1}^{T}(Y_{s,t} - \hat{\varphi}_s Z_{s,t})(Y_{s,t} - \hat{\varphi}_s Z_{s,t})',\ s = 1, \ldots, S. \tag{3.4}$$

These estimators are consistent under the hypothesis of cointegration as defined by (2.6). As a consequence, the continuous mapping theorem implies that substituting $(\hat{\Omega}_1, \ldots, \hat{\Omega}_S)$ in the log-likelihood (3.2) results in a function with identical (weak) convergence properties as the log-likelihood function evaluated for the maximum likelihood estimator, see Billingsley (1986). In other words,

$$\begin{aligned} & L(\varphi_1, \ldots, \varphi_S, \hat{\Omega}_1, \ldots, \hat{\Omega}_s) \\ = & -\frac{1}{2}T\sum_{s=1}^{S}\log|\hat{\Omega}_s| - \frac{1}{2}\sum_{s=1}^{S} tr\left(\sum_{t=1}^{T}\hat{\Omega}_s^{-1}\xi_{s,t}\xi'_{s,t}\right) \\ = & -\frac{1}{2}T\sum_{s=1}^{S}\log|\hat{\Omega}_s| - \frac{1}{2}\sum_{s=1}^{S} tr\left(\hat{\Omega}_s^{-1}\sum_{t=1}^{T}(Y_{s,t} - \hat{\varphi}_s Z_{s,t} + (\hat{\varphi}_s - \varphi_s)Z_{s,t})\right. \\ & \left.(Y_{s,t} - \hat{\varphi}_s Z_{s,t} + (\hat{\varphi}_s - \varphi_s)Z_{s,t})'\right) \\ = & -\frac{1}{2}T\sum_{s=1}^{S}\log|\hat{\Omega}_s| - \frac{1}{2}\sum_{s=1}^{S} tr\left(\hat{\Omega}_s^{-1}\sum_{t=1}^{T}(Y_{s,t} - \hat{\varphi}_s Z_{s,t})(Y_{s,t} - \hat{\varphi}_s Z_{s,t})'\right) \\ & -\frac{1}{2}\sum_{s=1}^{S} tr\left(\hat{\Omega}_s^{-1}\sum_{t=1}^{T}(\hat{\varphi}_s - \varphi_s)Z_{s,t}Z'_{s,t}(\hat{\varphi}_s - \varphi_s)'\right) \\ = & -\frac{1}{2}T\sum_{s=1}^{S}\log|\hat{\Omega}_s| - \frac{1}{2}SkT - \frac{1}{2}\sum_{s=1}^{S} tr\left(\hat{\Omega}_s^{-1}\sum_{t=1}^{T}(\hat{\varphi}_s - \varphi_s)Z_{s,t}Z'_{s,t}(\hat{\varphi}_s - \varphi_s)'\right). \end{aligned} \tag{3.5}$$

The first two elements of the final expression are independent of φ_s. Hence, when we optimize the log-likelihood given $(\hat{\Omega}_1, \ldots, \hat{\Omega}_s)$ over φ_s, $s = 1, \ldots, S$, we optimize only the last part.

Following Kleibergen (1998), we consider the relevant part of (3.5) as the objective function in a generalized method of moments approach. This objective function is now given by

$$\begin{aligned} G(\varphi_1, \ldots, \varphi_S) &= \sum_{s=1}^{S} tr\left(\hat{\Omega}_s^{-1} \sum_{t=1}^{T} (\hat{\varphi}_s - \varphi_s) Z_{s,t} Z_{s,t}' (\hat{\varphi}_s - \varphi_s)'\right) \\ &= \sum_{s=1}^{S} tr\left[\hat{\Omega}_s^{-1} \left(\sum_{t=1}^{T} Z_{s,t}\xi_{s,t}'\right)' \left(\sum_{t=1}^{T} Z_{s,t} Z_{s,t}'\right)^{-1} \left(\sum_{t=1}^{T} Z_{s,t}\xi_{s,t}'\right)\right]. \end{aligned} \tag{3.6}$$

This objective function can be considered as a quasi-log-likelihood since it has the same convergence properties as the log-likelihood. Notice that $G(\hat{\varphi}_1, \ldots, \hat{\varphi}_S)$ equals zero. This corresponds with no restrictions on φ_s. Taking all this together indicates that (3.6) can also be used as a quasi-likelihood ratio test statistic for cointegration restrictions which has the same convergence properties as the likelihood ratio statistic.

3.2. Generalized Method of Moments Estimators

We now turn to testing for cointegration in the PVAR(1) in (3.1). The objective function (3.6), however, does not directly reflect the parameters representing the hypothesis of cointegration. In order to obtain a specification of the objective function that contains these parameters in a straightforward way, we use that $\Pi_S + I_k = \alpha_S \beta_S' + I_k = \varphi_S \cdots \varphi_1$, to obtain

$$\begin{gathered} \xi_{s,t} = Y_{s,t} - \varphi_s Z_{s,t} \Leftrightarrow \\ \left(\varphi_S \cdots \varphi_{s+1} \xi_{s,t}\right) = \left(\varphi_S \cdots \varphi_{s+1}\right) Y_{s,t} - \left(\varphi_S \cdots \varphi_{s+1} \varphi_s \varphi_{s-1} \cdots \varphi_1\right) \left(\varphi_{s-1} \cdots \varphi_1\right)^{-1} Z_{s,t} \Leftrightarrow \\ \eta_{s,t} = V_{s,t} - \left(\varphi_S \cdots \varphi_1\right) W_{s,t} \Leftrightarrow \\ \eta_{s,t} = V_{s,t} - W_{s,t} - \alpha_S \beta_S' W_{s,t}, \end{gathered} \tag{3.7}$$

where we define $\eta_{s,t} \equiv (\varphi_S \cdots \varphi_{s+1})\xi_{s,t}$, $V_{s,t} \equiv \varphi_S \cdots \varphi_{s+1} Y_{s,t}$, and $W_{s,t} \equiv (\varphi_{s-1} \cdots \varphi_1)^{-1} Z_{s,t}$. Using (3.7), the objective function (3.6) can now be specified as

$$\begin{aligned} & G(\alpha_S, \beta_S | \varphi_1, \ldots, \varphi_S) \\ = & \sum_{s=1}^{S} tr\left[\Sigma_s^{-1} \left(\sum_{t=1}^{T} W_{s,t}\eta_{s,t}'\right)' \left(\sum_{t=1}^{T} W_{s,t} W_{s,t}'\right)^{-1} \left(\sum_{t=1}^{T} W_{s,t}\eta_{s,t}'\right)\right], \end{aligned} \tag{3.8}$$

where $\Sigma_s \equiv (\varphi_S \cdots \varphi_{s+1})\Omega_s(\varphi_S \cdots \varphi_{s+1})'$, $s = 1, \ldots, S$. Note that because of (2.9) we would obtain identical results when we would use another season instead of S. When we substitute $(\hat{\varphi}_1, \ldots, \hat{\varphi}_S, \hat{\Omega}_1, \ldots, \hat{\Omega}_S)$, see (3.3)-(3.4), in the objective function (3.8), we have

$$\begin{aligned} & G(\alpha_S, \beta_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S) \\ = & \sum_{s=1}^{S} tr\left[\hat{\Sigma}_s^{-1} \left(\sum_{t=1}^{T} \hat{W}_{s,t}\eta_{s,t}'\right)' \left(\sum_{t=1}^{T} \hat{W}_{s,t} \hat{W}_{s,t}'\right)^{-1} \left(\sum_{t=1}^{T} \hat{W}_{s,t}\eta_{s,t}'\right)\right], \end{aligned} \tag{3.9}$$

where $\hat{W}_{s,t} = (\hat{\varphi}_{s-1} \cdots \hat{\varphi}_1)^{-1} Z_{s,t}$ and $\hat{\Sigma}_s = (\hat{\varphi}_S \cdots \hat{\varphi}_{s+1})\hat{\Omega}_s(\hat{\varphi}_S \cdots \hat{\varphi}_{s+1})'$, and it is clear that this objective function now only depends on α_S and β_S.

To obtain the cointegrating vector estimator, that satisfies the first order condition for a minimum of (3.9), it is convenient to specify (3.9) as

$$\begin{pmatrix} vec(\sum_{t=1}^{T} \hat{W}_{1,t}\eta'_{1,t}\hat{\Sigma}_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} \hat{W}_{S,t}\eta'_{S,t}\hat{\Sigma}_S^{-\frac{1}{2}}) \end{pmatrix}' H^{-1} \begin{pmatrix} vec(\sum_{t=1}^{T} \hat{W}_{1,t}\eta'_{1,t}\hat{\Sigma}_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} \hat{W}_{S,t}\eta'_{S,t}\hat{\Sigma}_S^{-\frac{1}{2}}) \end{pmatrix}, \tag{3.10}$$

where $\hat{V}_{s,t} = \hat{\varphi}_S \cdots \hat{\varphi}_{s+1} Y_{s,t}$ and

$$H = \begin{pmatrix} I_k \otimes \sum_{t=1}^{T} \hat{W}_{1,t}\hat{W}'_{1,t} & 0 \cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 \cdots 0 & I_k \otimes \sum_{t=1}^{T} \hat{W}_{S,t}\hat{W}'_{S,t} \end{pmatrix}. \tag{3.11}$$

The cointegrating vector estimator for β_S then results from

$$\frac{\partial vec(G(\alpha_S, \beta_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S))}{\partial vec(\beta_S)'} = 0, \tag{3.12}$$

and it equals,

$$\begin{aligned} vec(\hat{\beta}_S) &= (D'H^{-1}D)D'H^{-1} \begin{pmatrix} vec(\sum_{t=1}^{T} \hat{W}_{1,t}(\hat{V}_{1,t} - \hat{W}_{1,t})'\hat{\Sigma}_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} \hat{W}_{S,t}(\hat{V}_{S,t} - \hat{W}_{S,t})'\hat{\Sigma}_S^{-\frac{1}{2}}) \end{pmatrix} \\ &= \left[\sum_{s=1}^{S} \left(\alpha'_S \hat{\Sigma}_s^{-1} \alpha_S \otimes \sum_{t=1}^{T} \hat{W}_{s,t}\hat{W}'_{s,t} \right) \right]^{-1} \\ &\quad \left[\sum_{s=1}^{S} vec \left(\sum_{t=1}^{T} \hat{W}_{s,t} \left(\hat{V}_{s,t} - \hat{W}_{s,t} \right)' \hat{\Sigma}_s^{-1} \alpha_S \right) \right], \end{aligned} \tag{3.13}$$

where

$$D = \begin{pmatrix} \Sigma_1^{-\frac{1}{2}} \alpha_S \otimes \sum_{t=1}^{T} \hat{W}_{1,t}\hat{W}'_{1,t} \\ \vdots \\ \Sigma_S^{-\frac{1}{2}} \alpha_S \otimes \sum_{t=1}^{T} \hat{W}_{S,t}\hat{W}'_{S,t} \end{pmatrix}, \tag{3.14}$$

Clearly, the cointegrating vector estimator (3.13) depends on the unknown parameter α_S. When we substitute a consistent estimator for α_S, the continuous mapping theorem implies that the limiting distribution of $\hat{\beta}_S$ is the same as that of the maximum likelihood estimator for all parameters jointly. The same holds for the limiting distribution of $G(\alpha_S, \beta_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S)$ which equals the limiting distribution of the corresponding likelihood ratio statistic. As we assume, $\beta_s = (I_r \; -\beta'_{2s})'$, consistent estimators of α_s, $s = 1, \ldots, S$, are the first r columns of $\hat{\Pi}_s$, see (2.5).

4. Limiting Distributions

To construct the limiting distribution of the cointegrating vector estimator (3.13) and the objective function (3.9), it is important to note that all $W_{s,t}$ defined in (3.7) have identical

limiting behavior for $s = 1, \ldots, S$. In fact, the limiting behavior of $W_{s,t}$ is characterized by

$$\frac{1}{\sqrt{T}} W_{s,t} \Rightarrow \frac{1}{\sqrt{T}} W_{S,t} \Rightarrow \beta_{S,\perp} \Phi_S^{\frac{1}{2}} B_{k-r}, \qquad s = 1, \ldots, S, \tag{4.1}$$

where

$$\Phi_S = (\alpha'_{S,\perp}\beta_{S,\perp})^{-1} \alpha'_{S,\perp} \left(\sum_{s=1}^{S} \Sigma_s \right) \alpha_{S,\perp} (\alpha'_{S,\perp}\beta_{S,\perp})^{-1\prime}. \tag{4.2}$$

We use these results to construct the limiting distributions of $\hat{\beta}_s$ and $G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S)$. Details of the relevant derivations are relegated to appendices A and B. We will now present only the final results.

Maximum likelihood estimators of cointegrating vectors in standard vector autoregressive models have mixed normal limiting distributions, see Johansen (1991) and Phillips (1991). Since we use only consistent estimators to construct the objective function (3.10), the limiting behavior of the objective function evaluated with the cointegrating vector estimator (3.13) is identical to the limiting behavior of the log-likelihood evaluated with the maximum likelihood estimator. The limiting distribution of the cointegrating vector estimator (3.13) is therefore identical to the limiting distribution of the maximum likelihood estimator of the cointegrating vector. In appendix A we show that the under the hypothesis $\Pi_S = \alpha_S \beta'_S$, it holds that

$$\begin{aligned} &T(vec(\hat{\beta}_{2S}) - vec(\beta_{2S})) \\ &\Rightarrow normal \left(0, \alpha_S \left(\textstyle\sum_{s=1}^{S} \Sigma_s^{-1} \right) \alpha_S^{-1} \otimes (\beta'_{S,\perp}\beta_{S,\perp})^{-1} \Phi_S^{-\frac{1}{2}\prime} \left(\int B_{k-r} B'_{k-r} \right)^{-1} \Phi_S^{-\frac{1}{2}} (\beta'_{S,\perp}\beta_{S,\perp})^{-1} \right), \end{aligned} \tag{4.3}$$

where $\int B_{k-r} B'_{k-r} = \int_0^1 B_{k-r}(t) B'_{k-r}(t) dt$, which is the notation we use throughout the whole paper. This result shows that the asymptotic normality of the maximum likelihood estimator of the cointegrating vector also holds in a PVAR(1) model.

In appendix B, we derive the limiting distribution of the likelihood ratio test for cointegration. Under the null hypothesis that $\Pi_S = \alpha_S \beta'_S$, this limiting distribution equals S times the limiting distribution of the Johansen trace statistic in the standard vector autoregressive model, that is

$$\begin{aligned} &G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S) \\ &\Rightarrow S \left\{ tr \left[\left(\int B_{k-r} dB'_{k-r} \right)' \left(\int B_{k-r} B'_{k-r} \right)^{-1} \left(\int B_{k-r} dB'_{k-r} \right) \right] \right\}. \end{aligned} \tag{4.4}$$

We can now test for cointegration by computing S^{-1} times $G(\hat{\alpha}_S, \hat{\beta}_S \,|\hat{\varphi}_1, \ldots, \hat{\varphi}_S)$ and compare it with the relevant critical values of the Johansen trace statistic, see Johansen (1995).

5. Model Extensions

The simple PVAR(1) allows us to focus only on the derivation of the results for cointegration without any further model complications. The use of the PVAR(1) model for practical purposes may be rather limited, however, as in practice often a more complicated lag structure and additional deterministic components are required, see Franses (1996) for example. In this section, we show how the above results generalize to extended PVAR models.

5.1. Deterministic Components

Deterministic components can be added to the PVAR (1) in (2.1) without altering much of the results from the previous sections. The most convenient specification to examine the impact of the deterministic components extends (3.1) and is

$$Y_{s,t} = c_s + d_s t + \varphi_s Z_{s,t} + \xi_{s,t}, \qquad s = 1, \ldots, S;\ t = 1, \ldots, T, \tag{5.1}$$

where c_s, d_s are $k \times 1$ of parameters. The representation in (5.1) allows us to use the same expressions of the estimators and the objective function as before, conditional on a well-known regression on the deterministic components, see Johansen (1995). Hence, instead of $Z_{s,t}$ and $Y_{s,t}$, we use

$$\begin{aligned} Z^*_{st} &= Z_{s,t} - \left[\sum_{t=1}^T Z_{s,t} \begin{pmatrix} 1 \\ t \end{pmatrix}'\right] \left[\sum_{t=1}^T \begin{pmatrix} 1 \\ t \end{pmatrix} \begin{pmatrix} 1 \\ t \end{pmatrix}'\right]^{-1} \begin{pmatrix} 1 \\ t \end{pmatrix} Z_{s,t}, \\ Y^*_{s,t} &= Y_{s,t} - \left[\sum_{t=1}^T Y_{s,t} \begin{pmatrix} 1 \\ t \end{pmatrix}'\right] \left[\sum_{t=1}^T \begin{pmatrix} 1 \\ t \end{pmatrix} \begin{pmatrix} 1 \\ t \end{pmatrix}'\right]^{-1} \begin{pmatrix} 1 \\ t \end{pmatrix} Y_{s,t}, \end{aligned} \tag{5.2}$$

in the expressions in section 3. In case only a constant term is present, the linear trend term is left out.

The limiting distribution of our objective function is affected by the deterministic components. In case of only a constant term, it becomes

$$\begin{aligned} &G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S) \Rightarrow \\ S\Bigg\{ tr\Bigg[&\left(\int \begin{pmatrix} B_{k-r-1} \\ \tau \end{pmatrix} dB'_{k-r})\right)' \left(\int \begin{pmatrix} B_{k-r-1} \\ \tau \end{pmatrix} \begin{pmatrix} B_{k-r-1} \\ \tau \end{pmatrix}'\right)^{-1} \\ &\left(\int \begin{pmatrix} B_{k-r-1} \\ \tau \end{pmatrix} dB'_{k-r})\right)\Bigg]\Bigg\}, \end{aligned} \tag{5.3}$$

where $\tau(t) = t$, $0 \le t \le 1$. In case of a constant and linear trend term,

$$\begin{aligned} &G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S) \Rightarrow \\ S\Bigg\{ tr\Bigg[&\left(\int \begin{pmatrix} B_{k-r-2} \\ \tau \\ \tau^2 \end{pmatrix} dB'_{k-r})\right)' \left(\int \begin{pmatrix} B_{k-r-2} \\ \tau \\ \tau^2 \end{pmatrix} \begin{pmatrix} B_{k-r-2} \\ \tau \\ \tau^2 \end{pmatrix}'\right)^{-1} \\ &\left(\int \begin{pmatrix} B_{k-r-2} \\ \tau \\ \tau^2 \end{pmatrix} dB'_{k-r})\right)\Bigg]\Bigg\}. \end{aligned} \tag{5.4}$$

Tables with critical values of the above limiting distributions (after division by S) are given in Johansen (1995).

The specification of the deterministic components in (5.1) assumes that they do not lie in the cointegration space. When we specify the deterministic components such that they do, as in, for example, Johansen (1991), the expressions of the estimators and objective functions become more complicated, and we consider these extensions as interesting further work.

5.2. Additional Lagged Variables

The results for the PVAR(1) model only extend to higher order dynamics, when we assume the following structure of a PVAR(p) model, that is

$$\begin{aligned} y_n &= \varphi_{1,s} y_{n-1} + \varphi_{2,s} y_{n-S-1} + \ldots + \varphi_{p,s} y_{n-(p-1)S-1} + u_n, \\ &= \varphi_s(L^S) y_{n-1} + u_n, \end{aligned} \tag{5.5}$$

where $\varphi_s(z) = \sum_{i=1}^{p} z^{i-1}\varphi_{i,s}$, with $\varphi_{i,s}$ $k \times k$ parameter matrices, $i = 1, \ldots, p$, $s = 1, \ldots, S$, and where p is the maximum lag length. This specification preserves the Markov structure that is needed to derive our earlier results.

The parameter matrix of (5.5), when written in stacked format, becomes

$$A(L) = \begin{pmatrix} I_k & 0 \cdots 0 & 0 \cdots 0 & -L\varphi_1(L) \\ -\varphi_2(L) & I_k & 0 \cdots 0 & 0 \\ 0 & 0 \quad \ddots & \ddots & 0 \\ 0 & 0 \cdots 0 & -\varphi_S(L) & I_k \end{pmatrix}, \tag{5.6}$$

where now L operates on annual data. The characteristic polynomial for (5.5) can therefore be written as

$$|A(z)| = |I_k - z\varphi_S(z) \cdots \varphi_1(z)| = 0, \tag{5.7}$$

where $\varphi_s(z)$, $s = 1, \ldots, S$, contains the parameters corresponding to season s. Cointegration in the PVAR (5.5) corresponds with $\Pi_S = \varphi_S(1) \cdots \varphi_1(1) - I_k$, having a lower rank. A convenient specification of the PVAR(p) model is therefore

$$\begin{aligned} y_n &= \varphi_s(L^S)y_{n-1} + u_n, \qquad s = 1, \ldots, S; \\ &= \varphi_s(1)y_{n-1} + \Delta_S\varphi_s^*(L^S)y_{n-1} + u_n, \\ &= \varphi_s(1)y_{n-1} + \varphi_{1,s}^*\Delta_S y_{n-1} + \ldots + \varphi_{p-1,s}^*\Delta_S y_{n-(p-2)S-1} + u_n, \end{aligned} \tag{5.8}$$

where $\Delta_S = 1 - L^S$, $\varphi_{i,s}^* = -\sum_{j=i+1}^{p} \varphi_{j,s}$, $i = 1, \ldots, p-1$, $s = 1, \ldots, S$. This specification allows us to use all the expressions of the cointegration estimators, objective functions and limiting distributions constructed previously by conducting a regression of $Z_{s,t}$, $Y_{s,t}$ on $\Delta Z_{s,t-1} \ldots \Delta Z_{s,t-p+2}$. When $Z_{s,t}^*$, $Y_{s,t}^*$ are the residuals of these regressions we can use them, instead of $Z_{s,t}$ and $Y_{s,t}$.

Appendix

A. Asymptotic normality of $\hat{\beta}_S$.

The cointegrating vector estimator reads,

$$vec(\hat{\beta}_S) = (D'H^{-1}D)D'H^{-1} \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}.$$

Under the null hypothesis, we have that $\Pi_S = \alpha_S\beta_S'$ and $V_{s,t} = W_{s,t} + \alpha_S\beta_S'W_{s,t} + \eta_{s,t}$ and hence $\hat{\beta}_S$ is characterized by,

$$\begin{aligned} vec(\hat{\beta}_S) &= (D'H^{-1}D)D'H^{-1} \\ &\quad \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(W_{1,t} + \alpha_S\beta_S'W_{1,t} + \eta_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(W_{S,t} + \alpha_S\beta_S'W_{S,t} + \eta_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\ &= (D'H^{-1}D)D'H^{-1} \left[Dvec(\beta_S + \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}\eta_{1,t}'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}\eta_{S,t}'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \right] \\ &= vec(\beta_S) + (D'H^{-1}D)D'H^{-1} \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}\eta_{1,t}'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}\eta_{S,t}'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}. \end{aligned}$$

The limiting behavior of $\hat{\beta}_S$ results from the limiting behavior of the different elements of the last term of $vec(\hat{\beta}_S)$.

The limiting behavior of D results from the limiting behavior of $W_{s,t}$, $s = 1, \ldots, S$, and can be characterized by:

$$\frac{1}{T^2}D \Rightarrow \begin{pmatrix} \Sigma_1^{-\frac{1}{2}}\alpha_S \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}' \\ \vdots \\ \Sigma_S^{-\frac{1}{2}}\alpha_S \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}' \end{pmatrix},$$

where $\int B_{k-r}B_{k-r}' = \int_0^1 B_{k-r}(t)B_{k-r}'(t)dt$. Similarly, the limiting behavior of H is obtained as

$$\frac{1}{T^2}H \Rightarrow \left(I_S \otimes \left(I_k \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}' \right) \right)$$

and thus

$$T^2H^{-1} \Rightarrow$$
$$\left(I_S \otimes \left(I_k \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime} \left(\int B_{k-r}B_{k-r}' \right)^{-1} \Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}' \right) \right),$$

and

$$\begin{aligned}
\tfrac{1}{T^2}D'H^{-1}D &\Rightarrow \left(\alpha_S'\left(\textstyle\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}}\textstyle\int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}'\right), \\
T^2(D'H^{-1}D)^{-1} &\Rightarrow \left(\left(\alpha_S'\left(\textstyle\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S\right)^{-1} \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}\right. \\
&\qquad \left.\left(\textstyle\int B_{k-r}B_{k-r}'\right)^{-1}\Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}'\right)
\end{aligned}$$

The limiting behavior of $\hat{\beta}_S$ is further characterized by,

$$\begin{aligned}
&\tfrac{1}{T}D'H^{-1}\begin{pmatrix} vec(\sum_{t=1}^{T}W_{1,t}\eta_{1,t}'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T}W_{S,t}\eta_{S,t}'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\
&\Rightarrow \left(\tfrac{1}{T^2}D'\right)(T^2H^{-1})\tfrac{1}{T}\begin{pmatrix} vec(\sum_{t=1}^{T}W_{1,t}\eta_{1,t}'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T}W_{S,t}\eta_{S,t}'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} \Sigma_1^{-\frac{1}{2}}\alpha_S\otimes I_k \\ \vdots \\ \Sigma_S^{-\frac{1}{2}}\alpha_S\otimes I_k \end{pmatrix}'\tfrac{1}{T}\begin{pmatrix} vec(\sum_{t=1}^{T}W_{1,t}\eta_{1,t}'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T}W_{S,t}\eta_{S,t}'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\
&\Rightarrow \tfrac{1}{T}\left[\textstyle\sum_{s=1}^{S}vec\left(\sum_{t=1}^{T}W_{s,t}\eta_{s,t}'\Sigma_s^{-1}\alpha_S\right)\right] \\
&\Rightarrow \tfrac{1}{T}\left\{vec\left[\textstyle\sum_{t=1}^{T}W_t\left(\sum_{s=1}^{S}\eta_{s,t}'\Sigma_s^{-1}\alpha_S\right)\right]\right\} \\
&\Rightarrow \tfrac{1}{T}\left\{vec\left[\textstyle\sum_{t=1}^{T}W_t\left(\sum_{s=1}^{S}\mu_{s,t}'\Sigma_s^{-\frac{1}{2}}\alpha_S\right)\right]\right\} \\
&\Rightarrow (I_k\otimes\beta_{S\perp}(\alpha_{S\perp}'\beta_{S\perp})^{-1})\left\{vec\left[\textstyle\int\alpha_{S\perp}'\left(\sum_{s=1}^{S}\Sigma_s\right)^{\frac{1}{2}}B_k dB_k'\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)^{\frac{1}{2}}\alpha_S\right]\right\} \\
&\Rightarrow vec\left\{\textstyle\int\beta_{S\perp}\Phi_S^{\frac{1}{2}}B_{k-r}dQ_r'\left[\alpha_S'\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S\right]^{\frac{1}{2}}\right\}
\end{aligned}$$

where $\mu_{s,t}$ equals $\Sigma_s^{-\frac{1}{2}}\eta_{s,t}$ which has an identity covariance matrix, and where Q_r and B_{k-r} are independent r and $k-r$ dimensional Brownian motion with identity covariance matrices. The latter independence follows from

$$\begin{aligned}
\beta_{S,\perp}\Phi_S^{\frac{1}{2}}B_{k-r} &= \beta_{Ss\perp}(\alpha_{S,\perp}'\beta_{S,\perp})^{-1}\alpha_{S,\perp}'\left(\sum_{s=1}^{S}\Sigma_s\right)^{\frac{1}{2}}B_k \\
&= \beta_{S,\perp}(\alpha_{S,\perp}'\beta_{S,\perp})^{-1}\alpha_{S,\perp}'(\iota\otimes I_k)'\begin{pmatrix}\Sigma_1 & 0\cdots 0 & 0\\ 0 & \ddots & 0\\ 0 & 0\cdots 0 & \Sigma_S\end{pmatrix}^{\frac{1}{2}}B_{k^2},
\end{aligned}$$

and

$$\begin{aligned}
\left[\alpha_S'\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S\right]^{\frac{1}{2}}Q_r &= \alpha_S'\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)^{\frac{1}{2}}B_k \\
&= \alpha_S'(\iota\otimes I_k)'\begin{pmatrix}\Sigma_1^{-1} & 0\cdots 0 & 0\\ 0 & \ddots & 0\\ 0 & 0\cdots 0 & \Sigma_S^{-1}\end{pmatrix}^{\frac{1}{2}}B_{k^2}.
\end{aligned}$$

where B_{k^2} is a k^2 dimensional Brownian motion with identity covariance matrix, and ι is an $S \times 1$ vector of ones. Furthermore,

$$\begin{aligned}
&\left[\alpha_S'\left(\iota\otimes I_k\right)'\begin{pmatrix}\Sigma_1^{-1} & 0\cdots 0 & 0\\ 0 & \ddots & 0\\ 0 & 0\cdots 0 & \Sigma_S^{-1}\end{pmatrix}^{\frac{1}{2}}\right]\left[\beta_{S,\perp}(\alpha_{S,\perp}'\beta_{S,\perp})^{-1}\alpha_{S,\perp}'\left(\iota\otimes I_k\right)\right.\\
&\left.\begin{pmatrix}\Sigma_1 & 0\cdots 0 & 0\\ 0 & \ddots & 0\\ 0 & 0\cdots 0 & \Sigma_S\end{pmatrix}^{\frac{1}{2}}\right]'\\
= \;& \alpha_S'\left(\iota\otimes I_k\right)'\left(\iota\otimes I_k\right)\alpha_{S,\perp}(\alpha_{S,\perp}'\beta_{S,\perp})^{-1\prime}\beta_{S,\perp}\\
= \;& S\alpha_S'\alpha_{S,\perp}(\alpha_{S,\perp}'\beta_{S,\perp})^{-1\prime}\beta_{S,\perp}=0,
\end{aligned}$$

as $\iota'\iota = S$. This shows that the scale matrices involved in the construction of Q_r and B_{k-r} from the k^2 dimensional Brownian motion B_{k^2} are orthogonal to one another. Q_r and B_{k-r} are therefore stochastically independent and the limiting distribution for $\hat{\beta}_S$ is mixed normal, see Phillips (1991).

Finally, the limiting distribution of $\hat{\beta}_S$ is given by

$$\begin{aligned}
& T(vec(\hat{\beta}_S)-vec(\beta_S))\\
\Rightarrow\;& \left[\left(\alpha_S'\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S\right)^{-1}\otimes\beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}\left(\int B_{k-r}B_{k-r}'\right)^{-1}\right.\\
& \left.\Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}'\right]vec\left[\int\beta_{S,\perp}\Phi_S^{-\frac{1}{2}}B_{k-r}dQ_r'\left(\alpha_S\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S\right)^{\frac{1}{2}}\right]\\
\Rightarrow\;& vec\left[\beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}\left(\int B_{k-r}B_{k-r}'\right)^{-1}\right.\\
& \left.\int B_{k-r}dQ_r'\left(\alpha_S\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S\right)^{-\frac{1}{2}}\right]\\
\Rightarrow\;& normal\left(0,\alpha_S\left(\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S^{-1}\right.\\
& \left.\otimes\beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}\left(\int B_{k-r}B_{k-r}'\right)^{-1}\Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}'\right).
\end{aligned}$$

This limiting distribution is degenerate because $\beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}\left(\int B_{k-r}B_{k-r}'\right)^{-1}\Phi_S^{-\frac{1}{2}}$ $(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}$ is a singular matrix. The limiting distribution of the unknown elements $\hat{\beta}_{2s}$ is, however, not degenerate. Because $\beta_{S\perp}=(\beta_{2s}\ I_{k-r})'$, we obtain

$$\begin{aligned}
& T(vec(\hat{\beta}_{2S})-vec(\beta_{2S}))\\
\Rightarrow normal&\left(0,\alpha_S\left(\textstyle\sum_{s=1}^{S}\Sigma_s^{-1}\right)\alpha_S^{-1}\otimes(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}\left(\int B_{k-r}B_{k-r}'\right)^{-1}\Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\right).
\end{aligned}$$

B. Convergence of the Objective Function

The objective function (3.10) evaluated in $(\hat{\alpha}_S, \hat{\beta}_S)$ is,

$$\begin{aligned} & G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \dots, \hat{\varphi}_S) \\ = & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}\hat{\eta}'_{1,t}\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}\hat{\eta}'_{S,t}\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}' H^{-1} \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}\hat{\eta}'_{1,t}\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}\hat{\eta}'_{S,t}\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}\hat{\eta}'_{1,t}\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}\hat{\eta}'_{S,t}\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\ = & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(W_{1,t} - W_{1,t} - \hat{\alpha}_S\hat{\beta}'_S W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t} - \hat{\alpha}_S\hat{\beta}'_S W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\ = & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} - Dvec\left(\hat{\beta}_S\right) \end{aligned}$$

$$= [(I_S \otimes (I_k \otimes I_k)) - D(D'H^{-1}D)^{-1}DH^{-1}] \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}.$$

This can be written as

$$\begin{aligned} & G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \dots, \hat{\varphi}_S) \\ = & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}' \left[H^{-1} - H^{-1}D(D'H^{-1}D)^{-1}D'H^{-1}\right] \\ & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \\ = & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}' D_\perp(D'_\perp H D_\perp)^{-1}D'_\perp \\ & \begin{pmatrix} vec(\sum_{t=1}^T W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^T W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}, \end{aligned}$$

as $H^{-1} - H^{-1}D(D'H^{-1}D)^{-1}D'H^{-1} = D_\perp(D_\perp' H D_\perp)^{-1}D_\perp'$, see Johansen (1991). The limiting behavior of $G(\hat{\alpha}_S, \hat{\beta}_S|\hat{\varphi}_1, \ldots, \hat{\varphi}_S)$ is determined by the limiting behavior of its elements under the hypothesis that $\Pi_S = \alpha_S \beta_S'$. We will discuss the limiting behavior of each of the elements.

The limiting behavior of D results from the limiting behavior of $W_{s,t}$ and can be characterized by

$$\frac{1}{T^2}D \Rightarrow \begin{pmatrix} \Sigma_1^{-\frac{1}{2}}\alpha_S \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}' \\ \vdots \\ \Sigma_S^{-\frac{1}{2}}\alpha_S \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}' \end{pmatrix}.$$

The limiting behavior of H is given by

$$\frac{1}{T^2}H \Rightarrow \left(I_S \otimes \left(I_k \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}'\right)\right).$$

Consequently,

$$T^2 D_\perp \Rightarrow \left(I_S \otimes \left(\Sigma_1^{\frac{1}{2}}\alpha_{S,\perp} \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}(\int B_{k-r}B_{k-r}')^{-1}\Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}'\right)\right)$$

as other specifications of $D_\perp$ have a different rate of convergence and disappear in the limit. So,

$$T^2 D_\perp' H D_\perp \Rightarrow$$
$$\left(I_S \otimes \left(\alpha_{S,\perp}'\Sigma_1\alpha_{S,\perp} \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}(\int B_{k-r}B_{k-r}')^{-1}\Phi_S^{-\frac{1}{2}}(\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}'\right)\right)$$

and therefore

$$\frac{1}{T^2}(D_\perp' H D_\perp)^{-1} \Rightarrow \left(I_S \otimes \left(\left(\alpha_{S,\perp}'\Sigma_1\alpha_{S,\perp}\right)^{-1} \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B_{k-r}'\Phi_S^{\frac{1}{2}\prime}\beta_{S,\perp}'\right)\right).$$

In a similar way, we obtain,

$$
\begin{aligned}
& T \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}' D_\perp \\
\Rightarrow \quad & \left[\left(T^2 D_\perp\right)' \frac{1}{T} \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix} \right]' \\
\Rightarrow \quad & \left[\left(I_S \otimes (I_k \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1} \Phi_S^{-\frac{1}{2}\prime} (\int B_{k-r}B_{k-r}')^{-1} \Phi_S^{-\frac{1}{2}} (\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}') \right) \right. \\
& \left. \frac{1}{T} \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(V_{1,t} - W_{1,t})'\alpha_{S,\perp}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(V_{S,t} - W_{S,t})'\alpha_{S,\perp}) \end{pmatrix} \right]' \\
\Rightarrow \quad & \left[\left(I_S \otimes (I_k \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1} \Phi_S^{-\frac{1}{2}\prime} (\int B_{k-r}B_{k-r}')^{-1} \Phi_S^{-\frac{1}{2}} (\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}') \right) \right. \\
& \left. \begin{pmatrix} vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB_k' \Sigma_1^{\frac{1}{2}} \alpha_{S\perp}) \\ \vdots \\ vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB_k' \Sigma_S^{\frac{1}{2}} \alpha_{S\perp}) \end{pmatrix} \right]' \\
\Rightarrow \quad & \left[\left(I_S \otimes (I_k \otimes \beta_{S,\perp}(\beta_{S,\perp}'\beta_{S,\perp})^{-1} \Phi_S^{-\frac{1}{2}\prime} (\int B_{k-r}B_{k-r}')^{-1} \Phi_S^{-\frac{1}{2}} (\beta_{S,\perp}'\beta_{S,\perp})^{-1}\beta_{S,\perp}') \right) \right. \\
& \left. \begin{pmatrix} vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB_{k-r}' (\alpha_{S,\perp}'\Sigma_1\alpha_{S,\perp})^{\frac{1}{2}}) \\ \vdots \\ vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB_{k-r}' (\alpha_{S,\perp}'\Sigma_S\alpha_{S,\perp})^{\frac{1}{2}}) \end{pmatrix} \right]'
\end{aligned}
$$

and thus that

$$
\begin{aligned}
& G(\hat{\alpha}_S, \hat{\beta}_S | \hat{\varphi}_1, \ldots, \hat{\varphi}_S) \\
= \quad & \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}' D_\perp (D_\perp' H D_\perp)^{-1} D_\perp' \\
& \begin{pmatrix} vec(\sum_{t=1}^{T} W_{1,t}(V_{1,t} - W_{1,t})'\Sigma_1^{-\frac{1}{2}}) \\ \vdots \\ vec(\sum_{t=1}^{T} W_{S,t}(V_{S,t} - W_{S,t})'\Sigma_S^{-\frac{1}{2}}) \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad & \begin{pmatrix} vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB'_{k-r}(\alpha'_{S,\perp}\Sigma_1\alpha_{S,\perp})^{\frac{1}{2}}) \\ \vdots \\ vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB'_{k-r}(\alpha'_{S,\perp}\Sigma_S\alpha_{S,\perp})^{\frac{1}{2}}) \end{pmatrix}' \\
& \left(I_S \otimes (I_k \otimes \beta_{S,\perp}(\beta'_{S,\perp}\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}(\int B_{k-r}B'_{k-r})^{-1}\Phi_S^{-\frac{1}{2}}(\beta'_{S,\perp}\beta_{S,\perp})^{-1}\beta'_{S,\perp})\right)' \\
& \left(I_S \otimes \left(\left(\alpha'_{S,\perp}\Sigma_1\alpha_{S,\perp}\right)^{-1} \otimes \beta_{S,\perp}\Phi_S^{\frac{1}{2}} \int B_{k-r}B'_{k-r}\Phi_S^{\frac{1}{2}\prime}\beta'_{S,\perp}\right)\right) \\
& \left(I_S \otimes (I_k \otimes \beta_{S,\perp}(\beta'_{S,\perp}\beta_{S,\perp})^{-1}\Phi_S^{-\frac{1}{2}\prime}(\int B_{k-r}B'_{k-r})^{-1}\Phi_S^{-\frac{1}{2}}(\beta'_{S,\perp}\beta_{S,\perp})^{-1}\beta'_{S,\perp})\right) \\
& \begin{pmatrix} vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB'_{k-r}(\alpha'_{S,\perp}\Sigma_1\alpha_{S,\perp})^{\frac{1}{2}}) \\ \vdots \\ vec(\int \beta_{S,\perp}\Phi_S^{\frac{1}{2}} B_{k-r} dB'_{k-r}(\alpha'_{S,\perp}\Sigma_S\alpha_{S,\perp})^{\frac{1}{2}}) \end{pmatrix} \\
\Rightarrow \quad & \begin{pmatrix} vec(\int B_{k-r}dB'_{k-r}) \\ \vdots \\ vec(\int B_{k-r}dB'_{k-r}) \end{pmatrix}' \left(I_S \otimes \left(I_{k-r} \otimes \left(\int B_{k-r}B'_{k-r}\right)^{-1}\right)\right) \begin{pmatrix} vec(\int B_{k-r}dB'_{k-r}) \\ \vdots \\ vec(\int B_{k-r}dB'_{k-r}) \end{pmatrix} \\
\Rightarrow \quad & S\left[tr\left(\left(\int B_{k-r}dB'_{k-r})\right)'\left(\int B_{k-r}B'_{k-r}\right)^{-1}\left(\int B_{k-r}dB'_{k-r})\right)\right],
\end{aligned}
$$

since all $W_{s,t}$ have the same limiting behavior.

References

[1] Billingsley, P. *Probability and Measure*. Wiley (New York), (1986).

[2] Boswijk, H.P., Franses, P.H. and N. Haldrup. Multiple Unit Roots in Periodic Autoregression. *Journal of Econometrics*, **80**:167–193, (1997).

[3] Engle, R.F. and C.W.J. Granger. Co-integration and error correction: Representation, estimation and testing. *Econometrica*, **55**:251–276, (1987).

[4] Franses, P.H. *Periodicity and Stochastic Trends in Economic Time Series*. Oxford University Press (Oxford), (1996).

[5] Hansen, L.P. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, **50**:1029–1054, (1982).

[6] Hylleberg, S., R.F. Engle, C.W.J. Granger and B.S. Yoo. Seasonal Integration and Cointegration. *Journal of Econometrics*, **44**:215–238, (1990).

[7] Johansen, S. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, **59**:1551–1580, (1991).

[8] Johansen, S. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford, (1995).

[9] Johansen, S. and E. Schaumburg. Likelihood Analysis of Seasonal Cointegration. *Journal of Econometrics*, **88**:301–339, (1998).

[10] Kleibergen, F. Reduced Rank Regression using GMM. In L. Matyas, editor, *Generalised Method of Moments Estimation*. Cambridge University Press, (1998).

[11] Lee, H.S. Maximum Likelihood Inference on Cointegration and Seasonal Cointegration. *Journal of Econometrics*, **54**:351–365, (1992).

[12] Lütkepohl, H. *Introduction to Multiple Time Series Analysis*. Springer Verlag (Berlin), (1991).

[13] Osborn, D.R. The Implications of Periodically Varying Coefficients for Seasonal Time-Series Processes. *Journal of Econometrics*, **48**:373–384, (1991).

[14] Phillips, P.C.B. Optimal Inference in Cointegrated Systems. *Econometrica*, **59**:283–306, (1991).

[15] Phillips, P.C.B. and S.N. Durlauf. Multiple Time Series Regression with Integrated Processes. *Review of Economic Studies*, **53**:473–495, (1986).

[16] Tiao, G.C. and M.R. Grupe. Hidden Periodic Autoregressive-Moving Average Models in Time Series Data. *Biometrika*, **67**:365–373, (1980).