Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models*

Jan J.J. Groen†
Tinbergen Institute, Erasmus University Rotterdam

Frank Kleibergen‡
Department of Quantitative Economics, University of Amsterdam

This version: July 28, 1999

Abstract

We propose in this paper a likelihood-based framework for cointegration analysis in panels of a fixed number of vector error correction models. Maximum likelihood estimators of the cointegrating vectors are constructed using iterated Generalized Method of Moments estimators. Using these estimators we construct likelihood ratio statistics to test for a common cointegration rank across the individual vector error correction models, both with heterogeneous and homogeneous cointegrating vectors. The corresponding limiting distributions are a summation of the limiting behavior of Johansen (1991) trace statistics. We also incorporate both unrestricted and restricted deterministic components which are either homogeneous or heterogeneous. The proposed framework is applied on a data set of exchange rates and appropriate monetary fundamentals. The test results show strong evidence for the validity of the monetary exchange rate model within a panel of vector error correction models for three major European countries, whereas the results based on individual vector error correction models for each of these countries separately are less supportive.

Keywords: likelihood, GMM, cointegration, panels of vector error correction models, common cointegration rank, exchange rates.
JEL classification: C12, C23, C51, F31, G15.

---

*This paper has benefited from helpful comments by seminar participants at the Tinbergen Institute Workshop on Time Series and Dynamic Panel Data in Amsterdam.

†Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands; E-mail: groen@few.eur.nl.

‡Department of Quantitative Economics, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands; E-mail: kleiberg@fee.uva.nl. Part of this research was conducted while Frank Kleibergen was affiliated with the Econometric Institute at the Erasmus University Rotterdam.

1 Introduction

Since the introduction of cointegration tests by Engle and Granger (1987), the usage of these tests on long-run relationships between non-stationary variables have grown in popularity amongst applied econometricians. A major fallacy of cointegration tests, however, is the need for a considerable span of the data. Shiller and Perron (1985) have shown that indeed the span of the data matters for the power of tests on non-stationary variables and not the number of observations. One way of dealing with this "time span" problem is to analyze time series across similar cross-sections in panel data sets. This approach has recently been applied on a large scale in testing the validity of purchasing power parity for exchange rates, as amongst others in Papell (1997) and O'Connell (1998). The panel data approach is the main focus of our paper.

One can distinguish two viewpoints in the literature on cointegration analysis within panels. The first approach can be typified as a panel version of the Engle and Granger (1987) residual-based two-step procedure. In that approach one estimates a long-run relationship on panel data and conduct Levin and Lin (1992)-type panel unit root tests on the residuals. This panel Engle-Granger procedure implies homogeneous long-run coefficients and the adjustment parameters, only the serial correlation in the residual panel unit root test is assumed to be heterogeneous. Variants of this approach can be found in Pedroni (1995) and Kao (1999), which differ in the way they deal with serial correlation. Groen (1999) has succesfully applied the panel Engle-Granger procedure on the monetary exchange rate model for 14 OECD countries. The main disadvantage of this approach is that differences in adjustment speeds and dynamics across the different individuals are not taken into account.

An opposite viewpoint is an approach in which all the model parameters and statistics are assumed to be heterogeneous in nature and independent of each other. In this case the panel cointegration tests are based on cross-sectional averages of the individual parameters and statistics.[1] Such an approach is only valid if one assumes that the individual model parameters and test statistics are determined independently of each other, and therefore does not use the panel dimension of the data. In our view the usage of the panel dimension is crucial in enhancing the power of cointegration testing, *i.e.* one should allow for interdependencies between the different individuals.

Our panel cointegration framework uses elements of both the pure time series-based cointegration approach and the pure panel data-based approaches as sketched above. We stack vector error correction (VEC) models of the different individuals into a joint panel VEC model. Within this panel VEC model we conduct cointegration rank tests on all the individual VEC models simultaneously, based on a common cointegration rank value. The corresponding likelihood ratio tests have limiting distributions which are based on a summation of the limiting behaviour of Johansen (1991, 1996) trace statistics across the individual VEC models. This is valid as long as one assumes a fixed cross-

[1]Like averages of residual Augmented Dickey-Fuller t-statistics in Pedroni (1995) and Kao (1999) or averages of likelihood ratio cointegration rank tests in Larsson *et al.* (1998).

section dimension. The accompanying canonical correlation-based maximum likelihood estimators, however, have unknown analytical expressions. We use iterated estimators based on the Generalized Method of Moments (GMM) framework of Hansen (1982) to construct maximum likelihood estimators of the cointegrating vectors. These iterative GMM estimators can be interpreted as analytical expressions of the appropriate maximum likelihood estimators, and these are used to construct likelihood ratio panel cointegration rank test statistics. The iterative GMM approach is based on the standard time series framework of Kleibergen (1999).

Our framework adds several novel features to the existing literature on panel cointegration. First, we use a maximum likelihood framework where we allow for an unrestricted disturbance covariance matrix within our panel. As such we allow for an instantaneous feedback between the different individuals in our panel. Contrary to this, existing studies assume an absence of cross-section correlation across the individual disturbances.[2] Assuming a (block) diagonal cross-section covariance structure can severely distort the size of test procedures when this assumption is inappropriate,[3] which at least is the case for exchange rate studies. A second novel feature is that the researcher is able to test *how many* cointegrating vectors the different individuals have in common within a panel. With respect to this only Larsson *et al.* (1998) has this feature, although they make use of the restrictive assumption of a block-diagonal cross-section covariance matrix. Finally, our framework can be used to test for the possibility of homogeneous long-run parameters combined with heterogeneous short-run dynamics.[4]

We proceed as follows. In section 2 we construct our panel VEC model by stacking the vector error correction models of each individual, and we show what the implications are of cointegration. We use in section 3 an iterated GMM framework to construct maximum likelihood estimators for both reduced rank and full rank panel VEC models, and we use these estimators to implement likelihood ratio panel cointegration testing. We also derive in this section the corresponding limiting distributions. Section 4 extends the panel VEC model to include higher order dynamics and several specifications of deterministic components. Our panel cointegration frame work is illustrated in section 5 through a test on the appropriateness of the monetary exchange rate model. Section 6 concludes.

In the remainder of this paper we use the following notations. The symbol "$\Rightarrow$" indicates weak convergence in probability measure. The trace of a matrix M is indicated with $tr(M)$ and vec(M) indicates vectorization of matrix M by stacking the columns of M. The integral $\int_0^1 B_j(t)dB_j'(t)dt$ is for short denoted as $\int B_j dB_j'$, where $B_j(t)$ is a j-dimensional vector Brownian motion with $0 \leq t \leq 1$ and an identity covariance matrix.

[2] An exception is Park and Ogaki (1991) who develop a seemingly unrelated regression analog for a system of canonical cointegrating regressions based on the approach of Park (1992). However, they only deal with cointegrating vector estimation and not with cointegration testing.

[3] This is shown in Monte Carlo experiments in O'Connell (1998) and Groen (1999) for the Levin and Lin (1992) panel unit root test and the panel Engle-Granger approach respectively.

[4] Pedroni (1996) based on the "Fully Modified OLS" approach of Phillips (1995), and Pesaran *et al.* (1998) based on single equation error correction models also has this feature, but they do not focus on cointegration testing.

2 The Panel Vector Error Correction Model

In this section we show how the vector error correction framework of Johansen (1991, 1996) can be adapted for use within dynamic panels. The panel vector error correction (VEC) model is build around standard VEC models for each of the individuals in the panel and we therefore start off with the standard time series VEC model.

In testing for cointegration between the k I(1) variables $z_{1t}, \ldots, z_{kt}$ for the standard time series case, we test the number of stationary linear relationships between these k variables. The standard time series framework for cointegration testing is to consider an unrestricted vector error correction model,[5]

$$\Delta y_t = \Pi y_{t-1} + \eta_{it}. \tag{1}$$

In (1) $\Delta y_t = y_t - y_{t-1}$, $y_t = (z_{1t} \cdots z_{kt})'$ and $\eta_t = (\eta_{1t} \cdots \eta_{kt})'$ are $k \times 1$ vectors. Note that the coefficient matrix Π has dimension $k \times k$, the disturbances η_t are distributed as $\eta_t \sim \mathrm{N}(0, \Sigma)$ and $t = 1, \ldots, T$. As in Johansen (1991, 1996), the number of cointegrating relationships for the k I(1) variables in y_t is tested using a reduced rank version of (1),

$$\Delta y_t = \alpha\beta' y_{t-1} + \eta_t. \tag{2}$$

Matrix α is the $k \times r$ matrix of adjustment coefficients and β is the $k \times r$ matrix of cointegrating vectors. The matrix of cointegrating vectors β can be normalized as,

$$\beta = \begin{pmatrix} I_r \\ -\beta_2 \end{pmatrix}, \tag{3}$$

where I_r is an $r \times r$ identity matrix and β_2 is a $(k-r) \times r$ matrix of unrestricted elements. The column dimension r of β indicates the cointegrating rank, *i.e.* the number stationary linear combinations of y_t.

The validity of reducing the rank of Π in (1) can be tested through likelihood ratio statistics $\mathrm{LR}(r|k)$ for each of the restrictions $r = 0, \ldots, k-1$ versus full rank k, where a full rank Π indicates that all k variables are I(0). In Johansen (1991, 1996) it is shown that the asymptotic distribution of $\mathrm{LR}(r|k)$ is a function of $(k-r)$ Brownian motions,

$$\mathrm{LR}(r|k) \Rightarrow tr\left(\int dB_{k-r}B'_{k-r}\left[\int B_{k-r}B'_{k-r}\right]^{-1}\int B_{k-r}dB'_{k-r}\right), \tag{4}$$

with B_{k-r} is a $(k-r)$-dimensional Brownian motion with an identity covariance matrix.

Within a panel of N individuals we are interested in cointegration testing with respect to the k I(1) variables in vector y_{it} of the i^{th} individual. In order to be able to do that,

[5]Higher order dynamics and deterministic components are not included at this stage. Section 4 deals with this issue.

we construct for each individual i a VEC model comparable with (1) and stack these into one system,

$$\begin{aligned}\Delta Y_t &= \begin{pmatrix} \Pi_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \Pi_N \end{pmatrix} Y_{t-1} + \varepsilon_t, \\ &= \Pi_{\mathrm{A}} Y_{t-1} + \varepsilon_t. \end{aligned} \tag{5}$$

The sub-matrices Π_i within the $Nk\times Nk$ matrix Π_{A} are of dimension $k\times k$ for $i = 1,\ldots,N$ and relates Δy_{it} to $y_{i,t-1}$. The panel VEC model in (5) consists of the $Nk \times 1$ vectors $Y_{t-1} = (y'_{1,t-1}\cdots y'_{N,t-1})'$, $\Delta Y_t = Y_t - Y_{t-1}$ and $\varepsilon_t = (\eta'_{1t}\cdots\eta'_{Nt})'$, with $t = 1,\ldots,T$. The disturbance vector ε_t contains the $k \times 1$ disturbance vectors η_{it} for each individual VEC model and $\varepsilon_t \sim \mathrm{N}(0,\ \Omega)$ with the $Nk \times Nk$ non-diagonal covariance matrix structure,

$$\Omega = \begin{pmatrix} \Omega_{11} & \cdots & \Omega_{1N} \\ \vdots & \ddots & \vdots \\ \Omega_{N1} & \cdots & \Omega_{NN} \end{pmatrix}. \tag{6}$$

The sub-matrix Ω_{ij} is of dimension $k \times k$ and $\Omega_{ij} \equiv \mathrm{Cov}(\eta_{it},\ \eta_{jt}) \neq 0$ for $i, j = 1,\ldots,N$.

The panel VEC model in (5) can be considered as a restricted version of the unrestricted full system VEC model,

$$\begin{aligned}\Delta Y_t &= \begin{pmatrix} \Pi_{11} & \cdots & \Pi_{1N} \\ \vdots & \ddots & \vdots \\ \Pi_{N1} & \cdots & \Pi_{NN} \end{pmatrix} Y_{t-1} + \varepsilon_t \\ &= \Pi_{\mathrm{ur}} Y_{t-1} + \varepsilon_t, \end{aligned} \tag{7}$$

which is a high dimensional version of VEC model (1) with Π_{ur} is $Nk \times Nk$. In (7) a $k \times k$ sub-matrix Π_{ij} relates Δy_{it} to $y_{j,t-1}$ for $i \neq j$ and Π_{ii} equals Π_i in (5). The block diagonal coefficient matrix Π_{A} in (5) implies a restriction on Π_{ur} which does not restrict the rank value and thus $\mathrm{rank}(\Pi_{\mathrm{A}}) = \mathrm{rank}(\Pi_{\mathrm{ur}}) = Nk$.

As we have $(Nk)^2$ parameters in Π_{ur} it is not efficient to estimate a VEC model like (7), even for moderate sizes of N and k, due to a large number parameters and the presence of spurious correlations. Next, we can observe that panel VEC model (5) imposes on VEC model (7) the restriction of *no* Granger causality between the different individuals in our panel. Toda and Phillips (1993) have shown that the limit distributions of test statistics for Granger causality within unrestricted VEC models depends on the rank of the true and *a priori* unknown long-run multiplier matrix. Therefore, as long as the true rank of Π_{ur} is unknown, test statistics for the hypothesis of zero values of the off-diagonal sub-matrices in Π_{ur} have an unknown asymptotic distribution. Hence, we base our panel cointegration analysis on the following assumption:

Assumption 2.1 *There is* ***no*** *linear dependence between the k I(1) variables of individual i and lags of the k I(1) variables of individual j for $i \neq j$, i.e. $\Pi_{ij} = 0$ for $i \neq j$ in (7).*

The block-diagonal structure of the long-run multiplier matrix implied by (5) is assumed to be valid ***a priori***.

Note that the non-diagonal disturbance covariance structure (6) does allow for *contemporaneous* dependence between the variables in y_{it} and y_{jt} for $i \neq j$, *i.e.* we allow for *instantaneous causality* in the sense of Lütkepohl (1993, Proposition 2.3).

Cointegration within our panel of N individuals now imposes rank reduction on the different Π_i's in (5). The imposed rank reduction is such that the cointegration rank is identical for all N individuals. We therefore obtain the following reduced rank specification of panel VEC model (5),

$$\begin{aligned}\Delta Y_t &= \begin{pmatrix} \alpha_1\beta_1' & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \alpha_N\beta_N' \end{pmatrix} Y_{t-1} + \varepsilon_t \\ &= \Pi_{\mathrm{B}} Y_{t-1} + \varepsilon_t. \end{aligned} \tag{8}$$

In (8) the matrices of adjustment parameters α_i and cointegrating vectors β_i are of dimension $k \times r$ for all the $i = 1, \ldots, N$ individuals. Cointegration testing within panel VEC model (5) is therefore based on the following *necessary* cross-section restriction:

Assumption 2.2 *When rank reduction of the Π_i's in (5) is appropriate, this rank reduction is* ***identical*** *for each individual i. Panel cointegration therefore implies a* ***common cointegration rank****: $rank(\Pi_i) = r$ for each $i = 1, \ldots, N$ and $r < k$. Panel cointegration rank tests are therefore identical to testing the restrictions implied by (8) on (5) and it reduces the rank value of the coefficient matrix Π_A in (5) from Nk to Nr.*

Next to the restriction of a common cointegration rank we can have the *optional* cross-section restriction of common cointegrating vectors,

$$\beta_i = \beta \quad \text{for} \quad i = 1, \ldots, N. \tag{9}$$

Both β_i and β are of dimension $k \times r$. Based on (9) we can rewrite (8) with common cointegrating vectors, *i.e.*

$$\begin{aligned}\Delta Y_t &= \begin{pmatrix} \alpha_1\beta' & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \alpha_N\beta' \end{pmatrix} Y_{t-1} + \varepsilon_t \\ &= \Pi_{\mathrm{C}} Y_{t-1} + \varepsilon_t. \end{aligned} \tag{10}$$

Johansen (1991, 1996) and Phillips (1991) derive that as long as the cointegration rank r is given, likelihood ratio tests of hypotheses on α and β within a pure time series VEC model like (2) have asymptotically a χ^2 distribution. As the restrictions in (9) do not alter the value of the common cointegration rank implied by panel VEC model (8), a likelihood ratio test of the restrictions in (9) also has a χ^2 distribution. Using normalization (3) for the individual cointegrating vectors, each β_i has $r(k-r)$ unrestricted elements, implying $(N-1)r(k-r)$ degrees of freedom in the χ^2 distribution. The likelihood ratio statistic $\mathrm{LR}(\Pi_{\mathrm{C}}|\Pi_{\mathrm{B}})$, which test the restrictions implied by (10) on (8), therefore has the following asymptotic distribution,

$$\mathrm{LR}(\Pi_{\mathrm{C}}|\Pi_{\mathrm{B}}) = 2[\ell_{\max}(\Pi_{\mathrm{B}},\Omega) - \ell_{\max}(\Pi_{\mathrm{C}},\Omega)] \Rightarrow \chi^2\left((N-1)r(k-r)\right). \tag{11}$$

In (11) $\ell_{\max}(\Pi_{\mathrm{B}},\Omega)$ and $\ell_{\max}(\Pi_{\mathrm{C}},\Omega)$ are the maximized log-likelihood functions corresponding with (8) and (10) respectively.

Using assumptions 2.1 and 2.2, panel cointegration testing in our framework is identical to testing the following null hypotheses versus the alternative hypothesis implied by panel VEC model (5):

$$\begin{aligned} &\mathrm{H}_0\colon \Pi_{\mathrm{B}} \text{ versus } \mathrm{H}_1\colon \Pi_{\mathrm{A}},\\ &\mathrm{H}_0\colon \Pi_{\mathrm{C}} \text{ versus } \mathrm{H}_1\colon \Pi_{\mathrm{A}}. \end{aligned} \tag{12}$$

The null hypotheses in (12) can be tested using likelihood ratio test statistics denoted with $\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}})$ and $\mathrm{LR}(\Pi_{\mathrm{C}}|\Pi_{\mathrm{A}})$ respectively. The panel VEC models in (5), (8) and (10) can be considered as composed of N standard time series k-variate VEC models. Hence, the asymptotic behaviour of $\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}})$ and $\mathrm{LR}(\Pi_{\mathrm{C}}|\Pi_{\mathrm{A}})$ can thus be based on the asymptotic behaviour of standard time series-based likelihood ratio cointegration rank tests as defined in (4). This result is conditional on the usage of large T asymptotics with cross-section dimension N being fixed. We deal with this issue in the next section.

3 Estimation and Testing of Panel Vector Error Correction Models

We construct in this section an iterative Generalized Method of Moments (GMM) framework to conduct likelihood ratio cointegration testing in panel VEC models and to construct maximum likelihood estimators of the cointegrating vectors. First, we establish in section 3.1 the link between maximum likelihood estimation and GMM-based estimation. We utilize GMM in section 3.2 to construct maximum likelihood estimators of the cointegrating vectors for our panel VEC models. In section 3.3 we use the iterated GMM estimates of both reduced rank and full rank panel VEC models to construct likelihood ratio test statistics for cointegration testing within our panel VEC models. We also derive in section 3.3 the corresponding asymptotic distributions.

3.1 GMM and Maximum Likelihood Estimation

The log-likelihood function for any of our panel VEC models can be specified as,

$$\ell(\Pi^*,\ \Omega) = -\frac{NkT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Omega| - \frac{1}{2}tr\left(\Omega^{-1}(\Delta Y - Y_{-1}\mathbf{\Pi})'(\Delta Y - Y_{-1}\mathbf{\Pi})\right), \tag{13}$$

where $\Pi^* = \Pi_{\mathrm{A}}$, Π_{B} or Π_{C}, $\mathbf{\Pi} = \Pi'_{\mathrm{A}}$, Π'_{B} or Π'_{C} and Ω has an identical structure as (6), all of which are defined in section 2. In (13) ΔY and Y_{-1} are $T \times Nk$ matrices

$$\Delta Y = \begin{pmatrix} \Delta Y'_1 \\ \vdots \\ \Delta Y'_T \end{pmatrix} \quad \text{and} \quad Y_{-1} = (Y_{1,-1} \cdots Y_{N,-1}) = \begin{pmatrix} Y'_0 \\ \vdots \\ Y'_{T-1} \end{pmatrix},$$

with the $T \times k$ matrix $Y_{i,-1}$ for $i = 1, \ldots, N$ and from section 2 Y_{t-1} is $Nk \times 1$ for $t = 1, \ldots, T$.

Log-likelihood function (13) can be rewritten as proportional to:

$$\begin{aligned} \ell(\Pi^*, \Omega) &\propto -\frac{T}{2}\ln|\Omega| - \frac{1}{2}tr\left(\Omega^{-1}(\Delta Y - Y_{-1}\mathbf{\Pi})'(\Delta Y - Y_{-1}\mathbf{\Pi})\right) \\ &\propto -\frac{T}{2}\ln|\Omega| - \frac{1}{2}\mathrm{vec}(\Delta Y - Y_{-1}\mathbf{\Pi})'(\Omega^{-1} \otimes I_T)\mathrm{vec}(\Delta Y - Y_{-1}\mathbf{\Pi}) \\ &\propto -\frac{T}{2}\ln|\Omega| - \frac{1}{2}\mathrm{vec}(\Delta Y - Y_{-1}\hat{\Pi}_{\mathrm{ur}})(\Omega^{-1} \otimes I_T)\mathrm{vec}(\Delta Y - Y_{-1}\hat{\Pi}_{\mathrm{ur}}) \\ &\qquad - \frac{1}{2}\mathrm{vec}(Y_{-1}(\hat{\Pi}_{\mathrm{ur}} - \mathbf{\Pi}))'(\Omega^{-1} \otimes I_T)\mathrm{vec}(Y_{-1}(\hat{\Pi}_{\mathrm{ur}} - \mathbf{\Pi})), \end{aligned} \tag{14}$$

with the $T \times T$ identity matrix I_T and the parameter estimate of VEC model (7) $\hat{\Pi}_{\mathrm{ur}} = (Y'_{-1}Y_{-1})^{-1}Y'_{-1}\Delta Y$. Note that in (14) we have expressed the log-likelihood for our panel VEC models in terms of the high-dimensional unrestricted VEC model (7). Using

$$\mathrm{vec}(\hat{\Pi}_{\mathrm{ur}}) = (\Omega^{-1} \otimes Y'_{-1}Y_{-1})^{-1}(\Omega^{-1} \otimes Y'_{-1})\mathrm{vec}(\Delta Y),$$

the third part of the last expression in (14) can be rewritten further as

$$\begin{aligned} &\mathrm{vec}(Y_{-1}(\hat{\Pi}_{\mathrm{ur}} - \mathbf{\Pi}))'(\Omega^{-1} \otimes I_T)\mathrm{vec}(Y_{-1}(\hat{\Pi}_{\mathrm{ur}} - \mathbf{\Pi})) \\ &\qquad = \mathrm{vec}(Y'_{-1}\boldsymbol{\varepsilon})'(\Omega \otimes Y'_{-1}Y_{-1})^{-1}\mathrm{vec}(Y'_{-1}\boldsymbol{\varepsilon}), \end{aligned} \tag{15}$$

where $\boldsymbol{\varepsilon} = \Delta Y - Y_{-1}\mathbf{\Pi}$.

Using (14), the conditional maximum likelihood estimator of Ω given Π^* equals:

$$\hat{\Omega}(\Pi^*) = \frac{1}{T}\left(\Delta Y - Y_{-1}\mathbf{\Pi}\right)'\left(\Delta Y - Y_{-1}\mathbf{\Pi}\right). \tag{16}$$

Expression (15) can be interpreted as a GMM objective function with all variables in Y_{-1} acting as instrument variables (see Kleibergen 1999):

$$G(\mathbf{\Pi},\ \Omega) = \mathrm{vec}(Y'_{-1}\boldsymbol{\varepsilon})'(\Omega \otimes Y'_{-1}Y_{-1})^{-1}\mathrm{vec}(Y'_{-1}\boldsymbol{\varepsilon}). \tag{17}$$

For the unrestricted system (7) the optimized objective function (17) is equal to zero, because the model is exactly identified in that case, *i.e.* the number of instruments and regressors are equal. Consequently, $\mathrm{vec}(Y'_{-1}\boldsymbol{\varepsilon})$ is then equal to zero. For the panel VEC models (5), (8) and (10), however, we have non-zero values for the optimized objective functions.

Based on (14) we can see that our GMM objective function is part of log-likelihood function (13),

$$\ell(\Pi^*,\ \Omega) = -\frac{NkT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Omega| - \frac{1}{2}tr\left(\Omega^{-1}\Delta Y' M_{Y_{-1}}\Delta Y\right) - \frac{1}{2}G(\mathbf{\Pi},\ \Omega), \tag{18}$$

where $M_{Y_{-1}} = I_T - Y_{-1}(Y'_{-1}Y_{-1})^{-1}Y'_{-1}$. In (18) we can estimate the disturbance covariance matrix Ω conditional on Π^* through estimator (16), and Π^* can be estimated given Ω using GMM objective function (17). Hence, we can maximize log-likelihood function (18) through sequentially applying the aforementioned estimation procedures of Ω and Π^* until convergence of the resulting estimators.

3.2 GMM-Based Estimation of Panel Vector Error Correction Models

In the unrestricted standard VEC model maximum likelihood estimators of the cointegrating vectors result from the canonical vectors associated with the canonical correlations as in Johansen (1991, 1996). The relationship between maximum likelihood and canonical correlations breaks down when we impose the restrictions implied by the panel VEC models (5), (8) and (10) on the unrestricted full-system VEC model (7) of section 2. Consequently, analytical expressions of the maximum likelihood estimator of the cointegrating vector cannot be obtained and one has to rely on numerical optimization. However, we can use GMM objective function (17) and covariance matrix estimator (16) to construct analytical expressions of the maximum likelihood cointegrating vector estimators.

Full Rank Estimation

GMM objective function (17) corresponding with panel VEC model (5) equals:

$$G(\mathbf{\Pi}_{\mathrm{A}}) = \mathrm{vec}(Y'_{-1}(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{A}}))'(\Omega \otimes Y'_{-1}Y_{-1})^{-1}\mathrm{vec}(Y'_{-1}(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{A}})), \tag{19}$$

where

$$\mathbf{\Pi}_{\mathrm{A}} = \Pi'_{\mathrm{A}} = \begin{pmatrix} \Pi'_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \Pi'_N \end{pmatrix}.$$

In (19) we use Y_{t-1} and not $y_{i,t-1}$ as instrument variables for each of the Nk equations of panel VEC model (5). The matrix of disturbances $\boldsymbol{\varepsilon} = \Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{A}}$ and the matrix of instrument variables across the Nk equations each must have identical structured cross-product matrices. If this is not the case we end up with unknown limit distributions of the GMM objective function and estimators. As we use Y_{t-1} as instruments for each equation in the panel VEC model, we can base our estimates and tests on a non-diagonal covariance matrix of the disturbances $\varepsilon_t = (\eta'_{1t} \cdots \eta'_{Nt})'$. Hence, in (19) the instruments Y_{-1} and the non-diagonal covariance matrix Ω correct for any contemporaneous correlation between the Nk I(1) variables in our panel.

We can specify in our objective function (19) the following:

$$\begin{aligned}\text{vec}(Y'_{-1}(\Delta Y - Y_{-1}\mathbf{\Pi}_{\text{A}})) = &\\ &\text{vec}\left(Y'_{-1}\Delta Y\right) - \text{vec}\left(Y'_{-1}Y_{1,-1}\Pi'_1 \ldots Y'_{-1}Y_{N,-1}\Pi'_N\right)\\ &= \text{vec}\left(Y'_{-1}\Delta Y\right) - F\begin{pmatrix}\text{vec}(\Pi'_1)\\ \vdots\\ \text{vec}(\Pi'_N)\end{pmatrix},\end{aligned} \tag{20}$$

where

$$F = \begin{pmatrix} (e_1 \otimes I_k) \otimes (Y'_{-1}Y_{1,-1}) & \cdots & (e_N \otimes I_k) \otimes (Y'_{-1}Y_{N,-1}) \end{pmatrix}. \tag{21}$$

In (21) e_i is the i^{th} N-dimensional unity vector and I_j is an $j \times j$ identity matrix, with j either equal to the number of variables k per individual VEC model or the total number of variables Nk. Objective function (19) can be minimized conditional on a consistent estimate of covariance matrix Ω for which we use the conditional maximum likelihood estimator $\hat{\Omega}(\Pi_{\text{A}})$, as defined in (16).

Minimizing objective function (19) with respect to $\Pi'_1, \ldots, \Pi'_N$ given $\hat{\Omega} = \hat{\Omega}(\Pi_{\text{A}})$ and (20) results in the following GMM-based estimator of Π_{A}, which equals estimation based on Seemingly Unrelated Regression Estimation (SURE), in panel VEC model (5):[6]

$$\begin{aligned}\begin{pmatrix}\text{vec}(\hat{\Pi}'_1)\\ \vdots\\ \text{vec}(\hat{\Pi}'_N)\end{pmatrix} &= (F'(\hat{\Omega} \otimes Y'_{-1}Y_{-1})^{-1}F)^{-1}F'(\hat{\Omega} \otimes Y'_{-1}Y_{-1})^{-1}\text{vec}(Y'_{-1}\Delta Y)\\ &= (F'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes Y'_{-1}Y_{-1})F_{\text{SURE}})^{-1}F'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes I_T)\text{vec}(Y'_{-1}\Delta Y),\end{aligned} \tag{22}$$

where

$$F = \left(I_{Nk} \otimes Y'_{-1}Y_{-1}\right) F_{\text{SURE}},$$

and

$$F_{\text{SURE}} = \begin{pmatrix} (e_1 \otimes I_k) \otimes (e_1 \otimes I_k) & \cdots & (e_N \otimes I_k) \otimes (e_N \otimes I_k) \end{pmatrix}. \tag{23}$$

To get maximum likelihood estimates of Π_{A} and the disturbance covariance matrix Ω we start off with a consistent *initial* estimate of Ω equal to

$$\hat{\Omega}(\hat{\Pi}_{\text{A,OLS}}) = \left(\hat{\Omega}_{ij}\right)_{i,j=1,\ldots,N} \text{ with } \hat{\Omega}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\hat{\eta}_{it}\hat{\eta}'_{jt}. \tag{24}$$

[6]The first order condition equals: $-F'(\hat{\Omega} \otimes Y'_{-1}Y_{-1})^{-1}[\text{vec}(Y'_{-1}\Delta Y) - F(\text{vec}(\Pi'_1) \cdots \text{vec}(\Pi'_N))'] = 0$.

In (24) $\hat{\eta}_{it}$ and $\hat{\eta}_{jt}$ are the OLS residuals from estimating the unrestricted VEC model in (1), see section 2, for each individual i and j separately with $i, j = 1, \ldots, N$. Using (24) we can estimate the Π_i's based on (22) and use these estimates to construct the following estimate of Ω:

$$\hat{\Omega}(\hat{\Pi}_{\mathrm{A}}) = \left(\hat{\Omega}_{ij}\right)_{i,j=1,\ldots,N} \text{ with } \hat{\Omega}_{ij} = \frac{1}{T}(\Delta Y_i - Y_{i,-1}\hat{\Pi}_i')'(\Delta Y_j - Y_{j,-1}\hat{\Pi}_j'). \tag{25}$$

Estimates of the Π_i's can now be constructed through (22) based on (25). We iterate this procedure until convergence of the estimators and the minimized GMM objective function. This iterative procedure yields maximum likelihood estimates of both the Π_i's and Ω, as we maximize likelihood function (18) with respect to *both* the Π_i's and Ω.

Heterogeneous Cointegrating Vectors

Panel VEC model (8) implies the following GMM objective function,

$$G(\mathbf{\Pi}_{\mathrm{B}}) = \mathrm{vec}(Y_{-1}'(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{B}}))'(\Omega \otimes Y_{-1}'Y_{-1})^{-1}\mathrm{vec}(Y_{-1}'(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{B}})), \tag{26}$$

where

$$\mathbf{\Pi}_{\mathrm{B}} = \Pi_{\mathrm{B}}' = \begin{pmatrix} \beta_1\alpha_1' & 0\cdots0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots0 & \beta_N\alpha_N' \end{pmatrix}. \tag{27}$$

In (27) β_i and α_i are $k \times r$ for $i = 1, \ldots, N$, where r equals the number of cointegrating vectors.

Within the objective function (26) we can write,

$$\begin{aligned} &\mathrm{vec}(Y_{-1}'(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{B}})) = \\ &\qquad \mathrm{vec}\left(Y_{-1}'\Delta Y\right) - \mathrm{vec}\left(Y_{-1}'Y_{1,-1}\beta_1\alpha_1' \ldots Y_{-1}'Y_{N,-1}\beta_N\alpha_N'\right) \\ &\qquad = \mathrm{vec}\left(Y_{-1}'\Delta Y\right) - E\begin{pmatrix} \mathrm{vec}(\beta_1) \\ \vdots \\ \mathrm{vec}(\beta_N) \end{pmatrix}, \end{aligned} \tag{28}$$

with

$$\begin{aligned} E &= \begin{pmatrix} (e_1 \otimes \alpha_1) \otimes (Y_{-1}'Y_{1,-1}) & \cdots & (e_N \otimes \alpha_N) \otimes (Y_{-1}'Y_{N,-1}) \end{pmatrix} \\ &= \left(I_{Nk} \otimes Y_{-1}'Y_{-1}\right) E_{\mathrm{SURE}}, \end{aligned} \tag{29}$$

using

$$E_{\mathrm{SURE}} = \begin{pmatrix} (e_1 \otimes \alpha_1) \otimes (e_1 \otimes I_k) & \cdots & (e_N \otimes \alpha_N) \otimes (e_N \otimes I_k) \end{pmatrix}.$$

The usage of selection matrix E_{SURE} in (29) indicates that we can interpret our GMM-estimator of $\beta_1, \ldots, \beta_N$ as a SURE-type estimator. Under the null hypothesis of a common cointegration rank r covariance matrix estimator (25) yields a consistent estimate of Ω for panel VEC model (8). Using normalization (3) from section 2 for $\beta_1, \ldots, \beta_N$, and under the true cointegration rank value r, a consistent estimator of α_i equals

$$\hat{\alpha}_i \equiv \text{the first } r \text{ columns of } \hat{\Pi}_i \text{ from (22) for } i = 1, \ldots, N. \tag{30}$$

Substituting (28), $\hat{\alpha}_1, \ldots, \hat{\alpha}_N$ and $\hat{\Omega} = \hat{\Omega}(\hat{\Pi}_{\mathrm{A}})$ in objective function (26), minimizing this objective function with respect to $\beta_1, \ldots, \beta_N$ result in the GMM-estimates

$$\begin{aligned} \begin{pmatrix} \mathrm{vec}(\hat{\beta}_1) \\ \vdots \\ \mathrm{vec}(\hat{\beta}_N) \end{pmatrix} &= (\hat{E}'(\hat{\Omega} \otimes Y_{-1}'Y_{-1})^{-1}\hat{E})^{-1}\hat{E}'(\hat{\Omega} \otimes Y_{-1}'Y_{-1})^{-1}\mathrm{vec}(Y_{-1}'\Delta Y) \\ &= (\hat{E}_{\mathrm{SURE}}'(\hat{\Omega}^{-1} \otimes Y_{-1}'Y_{-1})\hat{E}_{\mathrm{SURE}})^{-1}\hat{E}_{\mathrm{SURE}}'(\hat{\Omega}^{-1} \otimes I_T)\mathrm{vec}(Y_{-1}'\Delta Y), \end{aligned} \tag{31}$$

where $\hat{E}$ and $\hat{E}_{\mathrm{SURE}}$ equals (29) based on consistent estimates of the α_i's as defined in (30).

Given the estimate of the cointegrating vectors $\hat{\beta}_1, \ldots, \hat{\beta}_N$ and $\hat{\alpha}_1, \ldots, \hat{\alpha}_N$ from (30) we construct the estimate of $\mathbf{\Pi}_{\mathrm{B}}$ which we use in the conditional maximum likelihood estimator (16) of Ω to obtain $\hat{\Omega}(\hat{\Pi}_{\mathrm{B}})$. Jointly with $\hat{\beta}_1, \ldots, \hat{\beta}_N$, $\hat{\Omega}(\hat{\Pi}_{\mathrm{B}})$ is used to construct the GMM estimator of the α_i's. Minimizing GMM objective function (26) with respect to the α_i's conditional on the estimates $\hat{\beta}_i$ of β_i and $\hat{\Omega} = \hat{\Omega}(\hat{\Pi}_{\mathrm{B}})$ yields:

$$\begin{aligned} \begin{pmatrix} \mathrm{vec}(\hat{\alpha}_1') \\ \vdots \\ \mathrm{vec}(\hat{\alpha}_N') \end{pmatrix} &= (\Phi_{\mathrm{B}}'(\hat{\Omega} \otimes Y_{-1}'Y_{-1})^{-1}\Phi_{\mathrm{B}})^{-1}\Phi_{\mathrm{B}}'(\hat{\Omega} \otimes Y_{-1}'Y_{-1})^{-1}\mathrm{vec}(Y_{-1}'\Delta Y), \\ &= (\Phi_{\mathrm{B,SURE}}'(\hat{\Omega}^{-1} \otimes Y_{-1}'Y_{-1})\Phi_{\mathrm{B,SURE}})^{-1}\Phi_{\mathrm{B,SURE}}'(\hat{\Omega}^{-1} \otimes I_T)\mathrm{vec}(Y_{-1}'\Delta Y), \end{aligned} \tag{32}$$

where,

$$\begin{aligned} \Phi_{\mathrm{B}} &= \begin{pmatrix} (e_1 \otimes I_k) \otimes (Y_{-1}'Y_{1,-1}\hat{\beta}_1) & \cdots & (e_N \otimes I_k) \otimes (Y_{-1}'Y_{N,-1}\hat{\beta}_N) \end{pmatrix} \\ &= \left(I_{Nk} \otimes Y_{-1}'Y_{-1}\right)\Phi_{\mathrm{B,SURE}}, \end{aligned}$$

based on

$$\Phi_{\mathrm{B,SURE}} = \begin{pmatrix} (e_1 \otimes I_k) \otimes (e_1 \otimes \hat{\beta}_1) & \cdots & (e_N \otimes I_k) \otimes (e_1 \otimes \hat{\beta}_N) \end{pmatrix}.$$

In (32) Φ_{B} results from a similar respecification as in (28).

To obtain the maximum likelihood estimates of Ω, α_i and β_i for $i = 1, \ldots, N$, we sequentially apply estimators (31), $\hat{\Omega}(\hat{\Pi}_{\mathrm{B}})$ and (32) in an iterative way until convergence of the estimators. This iterative scheme can be outlined as follows:

0. Construct *initial* estimates of Ω and $\alpha_1, \ldots, \alpha_N$ through (25) and (30) respectively.

1. Construct the estimate of $\beta_1, \ldots, \beta_N$ through (31) given the estimates of $\alpha_1, \ldots, \alpha_N$ and Ω.

2. Construct estimator (16) for Ω given the estimated $\mathbf{\Pi}_{\mathrm{B}}$ which results from the estimates of $\alpha_1, \ldots, \alpha_N$ and $\beta_1, \ldots, \beta_N$ from step 1.

3. Construct estimator (32) for $\alpha_1, \ldots, \alpha_N$ given the estimates of $\beta_1, \ldots, \beta_N$ and $\hat{\Omega}(\hat{\Pi}_{\mathrm{B}})$ from steps 1 and 2.

4. When the objective function and estimators have not converged, go back to step 1.

The asymptotic behaviour of the maximum likelihood estimator of $\beta_1, \ldots, \beta_N$ that results from the above mentioned iterative scheme can be typified as follows:

Proposition 3.1 *Let,*

(a) *the cointegrating vectors β_i and loading factors α_i for $i = 1, \ldots, N$ not span orthogonal spaces, such that $\alpha'_{\perp i}\beta_{\perp i}$ is of full rank value,*

(b) *the estimators of α_i, β_i and Ω for $i = 1, \ldots, N$ be fully converged estimators that result from our iterative estimation scheme,*

(c) *the true common cointegration rank equal r,*

(d) *the cross-section dimension N be fixed and the time series dimension $T \to \infty$.*

Then the maximum likelihood cointegrating vector estimator based on (31) converges with rate T to its true value and its limiting distribution is a mixed normal distribution.

Proof: *See Appendix B.*

Homogeneous Cointegrating Vectors

The GMM objective function for panel VEC model (10), equals:

$$G(\mathbf{\Pi}_{\mathrm{C}}) = \mathrm{vec}(Y'_{-1}(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{C}}))'(\Omega \otimes Y'_{-1}Y_{-1})^{-1}\mathrm{vec}(Y'_{-1}(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{C}})), \qquad (33)$$

where

$$\mathbf{\Pi}_{\mathrm{C}} = \Pi'_{\mathrm{C}} = \begin{pmatrix} \beta\alpha'_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \beta\alpha'_N \end{pmatrix}.$$

In (33) we can rewrite $\mathrm{vec}(Y_{-1}'(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{C}}))$,

$$\begin{aligned}
&\mathrm{vec}(Y_{-1}'(\Delta Y - Y_{-1}\mathbf{\Pi}_{\mathrm{C}})) = \\
&\qquad \mathrm{vec}\left(Y_{-1}'\Delta Y\right) - \mathrm{vec}\left(Y_{-1}'Y_{1,-1}\beta\alpha_1' \ldots Y_{-1}'Y_{N,-1}\beta\alpha_N'\right) \\
&\qquad = \mathrm{vec}\left(Y_{-1}'\Delta Y\right) - \begin{pmatrix} \alpha_1 \otimes Y_{-1}'Y_{1,-1} \\ \vdots \\ \alpha_N \otimes Y_{-1}'Y_{N,-1} \end{pmatrix} \mathrm{vec}(\beta).
\end{aligned} \tag{34}$$

Based on (34) and the consistent estimates $\hat{\alpha}_1, \ldots, \hat{\alpha}_N$ and $\hat{\Omega}$, as defined in (30) and (25), minimization of (33) with respect to β yields the following estimate of β:

$$\begin{aligned}
\mathrm{vec}(\hat{\beta}) &= \left(\hat{D}'\left(\hat{\Omega} \otimes Y_{-1}'Y_{-1}\right)^{-1}\hat{D}\right)^{-1}\hat{D}'\left(\hat{\Omega} \otimes Y_{-1}'Y_{-1}\right)^{-1}\mathrm{vec}\left(Y_{-1}'\Delta Y\right) \\
&= \left(\hat{D}_{\mathrm{SURE}}'\left(\hat{\Omega}^{-1} \otimes Y_{-1}'Y_{-1}\right)\hat{D}_{\mathrm{SURE}}\right)^{-1}\hat{D}_{\mathrm{SURE}}'\left(\hat{\Omega}^{-1} \otimes I_T\right)\mathrm{vec}\left(Y_{-1}'\Delta Y\right),
\end{aligned} \tag{35}$$

where $\hat{D}$ and $\hat{D}_{\mathrm{SURE}}$ equal

$$\begin{aligned}
D &= \begin{pmatrix} \alpha_1 \otimes Y_{-1}'Y_{1,-1} \\ \vdots \\ \alpha_N \otimes Y_{-1}'Y_{N,-1} \end{pmatrix} = \left(I_{Nk} \otimes Y_{-1}'Y_{-1}\right)D_{\mathrm{SURE}}, \\
D_{\mathrm{SURE}} &= \begin{pmatrix} \alpha_1 \otimes (e_1 \otimes I_k) \\ \vdots \\ \alpha_N \otimes (e_N \otimes I_k) \end{pmatrix}
\end{aligned} \tag{36}$$

in which we substitute estimates of $\alpha_1, \ldots, \alpha_N$. Given the estimates of $\alpha_1, \ldots, \alpha_N$ and β, through (30) and (35), we can then construct the estimated $\mathbf{\Pi}_{\mathrm{C}}$ which we use in estimator (16) for Ω.

Conditional on the estimates $\hat{\beta}$ and $\hat{\Omega} = \hat{\Omega}(\hat{\Pi}_{\mathrm{C}})$, minimization of objective function (33) with respect to $\alpha_1, \ldots, \alpha_N$ yields a GMM estimator of the α_i's:

$$\begin{aligned}
\begin{pmatrix} \mathrm{vec}(\hat{\alpha}_1') \\ \vdots \\ \mathrm{vec}(\hat{\alpha}_N') \end{pmatrix} &= (\Phi_{\mathrm{C}}'(\hat{\Omega} \otimes Y_{-1}'Y_{-1})^{-1}\Phi_{\mathrm{C}})^{-1}\Phi_{\mathrm{C}}'(\hat{\Omega} \otimes Y_{-1}'Y_{-1})^{-1}\mathrm{vec}(Y_{-1}'\Delta Y) \\
&= (\Phi_{\mathrm{C,SURE}}'(\hat{\Omega}^{-1} \otimes Y_{-1}'Y_{-1})\Phi_{\mathrm{C,SURE}})^{-1}\Phi_{\mathrm{C,SURE}}'(\hat{\Omega}^{-1} \otimes I_T)\mathrm{vec}(Y_{-1}'\Delta Y),
\end{aligned} \tag{37}$$

where,

$$\begin{aligned}
\Phi_{\mathrm{C,SURE}} &= \left((e_1 \otimes I_k) \otimes (e_1 \otimes \hat{\beta}) \quad \cdots \quad (e_N \otimes I_k) \otimes (e_N \otimes \hat{\beta}) \right), \\
\Phi_{\mathrm{C}} &= \left(I_{Nk} \otimes Y_{-1}'Y_{-1}\right)\Phi_{\mathrm{C,SURE}} \\
&= \left((e_1 \otimes I_k) \otimes (Y_{-1}'Y_{1,-1}\hat{\beta}) \quad \cdots \quad (e_N \otimes I_k) \otimes (Y_{-1}'Y_{N,-1}\hat{\beta}) \right).
\end{aligned}$$

Maximum likelihood estimation of β, Ω and $\alpha_1, \dots, \alpha_N$ can be done in an analogous way as in the case of heterogeneous cointegrating vectors through (iterative) sequential estimation of β, Ω and the α_i's based on (35), (16) and (37) respectively.

3.3 Likelihood Ratio Testing

The maximized value of log-likelihood function (18) in section 3.1 can be obtained by substituting the optimized GMM objective functions and the conditional maximum likelihood estimates of Ω in (18). Likelihood ratio cointegration rank testing within our panel VEC models can then be conducted based on these maximized log-likelihood functions.

The discussion in section 3.1 indicates that we can write, given estimate $\hat{\Pi}^*$ of Π^*, in (18)

$$tr\left(\hat{\Omega}(\hat{\Pi}^*)^{-1}\Delta Y' M_{Y_{-1}}\Delta Y\right) + G(\mathbf{\Pi},\ \hat{\Omega}(\hat{\Pi}^*)) = tr\left(\hat{\Omega}(\hat{\Pi}^*)^{-1}(\Delta Y - Y_{-1}\mathbf{\Pi})'(\Delta Y - Y_{-1}\mathbf{\Pi})\right),$$

and based on this result it is straightforward to show that the maximized value of (18) given Π^* equals:

$$\ell_{\max}(\hat{\Pi}^*,\ \hat{\Omega}(\hat{\Pi}^*)) = \text{constant} - \frac{T}{2}\ln|\hat{\Omega}(\hat{\Pi}^*)|, \tag{38}$$

with $\hat{\Pi}^* = \hat{\Pi}_{\mathrm{A}}$, $\hat{\Pi}_{\mathrm{B}}$ or $\hat{\Pi}_{\mathrm{C}}$. The maximized log-likelihood functions, as defined in (38), can be used to conduct likelihood ratio testing of the coefficient matrices implied by the panel VEC models (8) and (10) versus the coefficient matrix Π_{A} in (5), as summarized under (12) in section 2:

$$\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) = 2[\ell(\hat{\Pi}_{\mathrm{A}},\hat{\Omega}(\hat{\Pi}_{\mathrm{A}})) - \ell(\hat{\Pi}_{\mathrm{B}},\hat{\Omega}(\hat{\Pi}_{\mathrm{B}}))] = T[\ln|\hat{\Omega}(\hat{\Pi}_{\mathrm{B}})| - \ln|\hat{\Omega}(\hat{\Pi}_{\mathrm{A}})|], \tag{39}$$

and

$$\mathrm{LR}(\Pi_{\mathrm{C}}|\Pi_{\mathrm{A}}) = 2[\ell(\hat{\Pi}_{\mathrm{A}},\hat{\Omega}(\hat{\Pi}_{\mathrm{A}})) - \ell(\hat{\Pi}_{\mathrm{C}},\hat{\Omega}(\hat{\Pi}_{\mathrm{C}}))] = T[\ln|\hat{\Omega}(\hat{\Pi}_{\mathrm{C}})| - \ln|\hat{\Omega}(\hat{\Pi}_{\mathrm{A}})|]. \tag{40}$$

In (39) and (40) $\hat{\Omega}(\hat{\Pi}_{\mathrm{A}})$, $\hat{\Omega}(\hat{\Pi}_{\mathrm{B}})$ and $\hat{\Omega}(\hat{\Pi}_{\mathrm{C}})$ result from the iterative estimation procedures of section 3.2.

The main implication of the restrictions implied by panel VEC model (8) on (5) in section 2 is that we test for identical rank reduction in N individual VEC models simultaneously.[7] Therefore, based on $T \to \infty$ one can motivate the limiting distribution of $\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}})$ as the sum of $(k-r)$-dimensional Brownian motion functionals, as defined in (4) in section 2, across N individuals. This line of reasoning is valid as long as one uses a maximum likelihood estimation procedure which takes into account the correlations across all Nk variables in our panel. The asymptotic behaviour of the LR-statistic in (39) can thus be defined as:

[7] This is also indicated by the representation theorem of panel VEC model (8) in Appendix A.

Proposition 3.2 *Given the conditions from proposition 3.1, the limiting distribution of* $LR(\Pi_B|\Pi_A)$ *in (39) equals:*

$$LR(\Pi_B|\Pi_A) \Rightarrow \sum_{i=1}^{N} tr\left(\int dB_{k-r,i}B'_{k-r,i}\left[\int B_{k-r,i}B'_{k-r,i}\right]^{-1}\int B_{k-r,i}dB'_{k-r,i}\right). \tag{41}$$

In (41) $B_{k-r,i}$ *is a* $(k-r)$*-dimensional Brownian motion for individual* i *with an identity covariance matrix.*
Proof: *See Appendix C.*

The likelihood ratio statistic $\mathrm{LR}(\Pi_\mathrm{C}|\Pi_\mathrm{A})$ for testing the restrictions implied by panel VEC model (10) on panel VEC model (5) can be decomposed as follows:

$$\mathrm{LR}(\Pi_\mathrm{C}|\Pi_\mathrm{A}) = \mathrm{LR}(\Pi_\mathrm{C}|\Pi_\mathrm{B}) + \mathrm{LR}(\Pi_\mathrm{B}|\Pi_\mathrm{A}). \tag{42}$$

In (42) the asymptotic distribution corresponding with $\mathrm{LR}(\Pi_\mathrm{B}|\Pi_\mathrm{A})$ equals (41) in proposition 3.2. The conditional likelihood ratio statistic $\mathrm{LR}(\Pi_\mathrm{C}|\Pi_\mathrm{B})$ which tests the restriction of homogeneous cointegrating vectors given the common cointegration rank has a limiting distribution as defined in (11) in section 2. We can therefore define the limiting behaviour of LR-statistic (40) as:

Proposition 3.3 *Given the conditions from proposition 3.1, the limiting distribution of* $LR(\Pi_C|\Pi_A)$ *in (40) equals:*

$$LR(\Pi_C|\Pi_A) \Rightarrow$$
$$\chi^2\left((N-1)r(k-r)\right) + \sum_{i=1}^{N} tr\left(\int dB_{k-r,i}B'_{k-r,i}\left[\int B_{k-r,i}B'_{k-r,i}\right]^{-1}\int B_{k-r,i}dB'_{k-r,i}\right). \tag{43}$$

Proof: *This result follows from decomposing* $LR(\Pi_C|\Pi_A)$ *as in (42). We normalize the cointegrating vectors as in (3):* $\beta'_i = (I_r \;\; -\beta'_{2i})$*. In Appendix B we show that the maximum likelihood cointegrating vector estimator of the* β_{2i}*'s based on (31) is asymptotically distributed as a mixed normal distribution. The assumption of homogeneous cointegrating vectors restricts the values of the* $r(k-r)$ *elements in the* β_{2i} *parameter matrix for* $(N-1)$ *individuals. Hence, the quasi-likelihood ratio test statistic* $LR(\Pi_C|\Pi_B)$ *tests a restriction on* $(N-1)r(k-r)$ *parameters and this statistic is asymptotically distributed as a* $\chi^2((N-1)r(k-r))$ *random variable. The limiting distribution of* $LR(\Pi_B|\Pi_A)$ *results directly from proposition 3.2.*

4 Deterministic Components and Higher Order Dynamics

The panel VEC models as defined in (5), (8), and (10) have no higher order dynamics and deterministic components. In practice, however, VEC models contain these components and in this section we discuss how the analysis changes when we include them. We first discuss in section 4.1 the issue of the deterministic components. In section 4.2 the estimation of higher order dynamics is discussed.

4.1 Deterministic Components

Several specifications of the deterministic components are possible and each of them influence the limiting distribution of the LR statistics of section 3.3. Consider the VEC model of individual i,

$$\Delta y_{it} = \alpha_i \beta_i' y_{i,t-1} + \delta_i' x_t + \varepsilon_{it}, \tag{44}$$

where δ_i is $m \times k$ and $i = 1, \dots, N$. The vector x_t is $m \times 1$ and it contains the deterministic components that are identical across the individuals. Using (44) we can distinguish several different specifications of the deterministic components:

B.1. Heterogeneous cointegrating vectors and unrestricted deterministic components, *i.e.* in (44) β_i and δ_i are unrestricted for $i = 1, \dots, N$.

B.2. Heterogeneous cointegrating vectors and the deterministic components are restricted to lie in the cointegration space, *i.e.* in (44) β_i is unrestricted for $i = 1, \dots, N$ and $\delta_i = \mu_i \alpha_i'$ with μ_i is $m \times r$.

C.1. Homogeneous cointegrating vectors and unrestricted deterministic components, *i.e.* in (44) $\beta_i = \beta$ for $i = 1, \dots, N$ and δ_i is unrestricted.

C.2. Homogeneous cointegrating vectors and the deterministic components are heterogeneous and restricted to lie in the cointegration space, *i.e.* in (44) $\beta_i = \beta$ for $i = 1, \dots, N$ and $\delta_i = \mu_i \alpha_i'$, with μ_i is $m \times r$.

C.3. Homogeneous cointegrating vectors and the deterministic components are homogeneous and restricted to lie in the cointegration space, *i.e.* in (44) $\beta_i = \beta$ and $\delta_i = \mu \alpha_i'$ for $i = 1, \dots, N$, with μ is $m \times r$.

We only briefly discuss the consequences of specifications **B.1-C.3** for the GMM-estimators and the limiting distributions of the LR statistics. This discussion is based on the discussion in sections 3.2 and 3.3. Only a constant is used in our discussion of the deterministic components as other cases follow straightforwardly, *i.e.* in (44) we have the $k \times 1$ vector $\delta_i' = c_i$ for the unrestricted case and for the restricted case μ_i and μ are $1 \times r$.

Including constants in the full rank panel VEC model (5) from section 2, implies that these constants are always unrestricted and the corresponding $(Nk + 1) \times Nk$ coefficient matrix equals:

$$\mathbf{\Pi}_{\text{A.1}} = \Pi_{\text{A.1}}' = \begin{pmatrix} \Pi_1' & 0 \cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 \cdots 0 & \Pi_N' \\ c_1' & \cdots & c_N' \end{pmatrix}. \tag{45}$$

Analogous to Johansen (1991, 1996), we can concentrate out these unrestricted constants through OLS regressions of Δy_{it} and $y_{i,t-1}$ on a constant. Therefore, we can define

$$\begin{aligned}\Delta\tilde{Y}_i = M_\iota \Delta Y_i \text{ and } \Delta\tilde{Y} = (\Delta\tilde{Y}_1 \cdots \Delta\tilde{Y}_N),\\ \tilde{Y}_{i,-1} = M_\iota Y_{i,-1} \text{ and } \tilde{Y}_{-1} = (\tilde{Y}_{1,t-1} \cdots \tilde{Y}_{N,-1}),\end{aligned} \tag{46}$$

where $M_\iota = I_T - \iota(\iota'\iota)^{-1}\iota'$ with the $T \times 1$ vector of ones ι. Hence, the minimized GMM objective function $G(\Pi_{\text{A.1}},\ \hat{\Omega})$ equals (19) from section 3.2 based on $\Pi_1, \ldots, \Pi_N$ from (45) and the variables in (46). The estimator of the Π_i's in (45) now equals estimator (22) from section 3.2 with $\Delta\tilde{Y}_i$, $\tilde{Y}_{i,-1}$ and $\tilde{Y}_{-1}$.

In the remainder of this subsection, we construct the GMM cointegrating vector estimator using consistent estimates of $\hat{\alpha}_1, \ldots, \hat{\alpha}_N$ and $\hat{\Omega}$ from (30) and (25) based on (46). These estimators can then be used in an iterative scheme jointly with estimators for Ω and $\alpha_1, \ldots, \alpha_N$ based on (16) and (32) or (37) respectively, as in section 3.2. Upon convergence of the estimators and the objective function, the resulting estimates can then be considered as maximum likelihood estimates. We only discus the GMM estimators for the cointegrating vector as the GMM estimators for $\alpha_1, \ldots, \alpha_N$ are identical to those of section 3.2.

Unrestricted constants: specifications B.1 and C.1

For specification B.1, the corresponding panel VEC model is similar to (8) from section 2 based on the $(Nk+1) \times Nk$ parameter matrix

$$\mathbf{\Pi}_{\text{B.1}} = \Pi'_{\text{B.1}} = \begin{pmatrix} \beta_1\alpha'_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \beta_N\alpha'_N \\ c'_1 & \cdots & c'_N \end{pmatrix}. \tag{47}$$

The optimal estimates of $\beta_1, \ldots, \beta_N$ in (47) follow from a GMM objective function $G(\mathbf{\Pi}_{\text{B.1}},\ \hat{\Omega})$ which is identical to (26) from section 3.2 based on the variables in (46). The estimator for the heterogeneous cointegrating vectors is now identical to (31) from section 3.2 where ΔY, $Y_{i,-1}$ and Y_{-1} are replaced with $\Delta\tilde{Y}$, $\tilde{Y}_{i,-1}$ and $\tilde{Y}_{-1}$ from (46). The corresponding LR cointegration rank test can be calculated in an analogous way as in section 3.3, based on iterative estimation.

In a similar way specification C.1 implies a long-run multiplier matrix

$$\mathbf{\Pi}_{\text{C.1}} = \Pi'_{\text{C.1}} = \begin{pmatrix} \beta\alpha'_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \beta\alpha'_N \\ c'_1 & \cdots & c'_N \end{pmatrix}. \tag{48}$$

The appropriate GMM objective function $G(\Pi_{\text{C.1}},\ \hat{\Omega})$ is based on (33) from section 3.2 with the variables from (46). GMM estimates of the homogeneous cointegrating vectors in (48) are equal to (35) from section 3.2 based on the variables of (46).

Restricted constants: specifications B.2, C.2 and C.3

The GMM objective function with restricted constants, resulting from adjusting the log-likelihood function from section 3.1 for this case, is similar to:

$$G(\mathbf{\Pi},\ \hat{\Omega}) = \text{vec}(Z'_{-1}(\Delta Y - Z_{-1}\mathbf{\Pi})'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\text{vec}(Z'_{-1}(\Delta Y - Z_{-1}\mathbf{\Pi})), \tag{49}$$

with $Z_{-1} = (Y_{-1} \ \ \iota)$, the $T \times 1$ vector of ones ι and the $(Nk + 1) \times Nk$ matrix $\mathbf{\Pi}$.

If we have specification B.2, the corresponding GMM objective function $G(\mathbf{\Pi}_{\text{B.2}},\ \hat{\Omega})$ equals (49) with $\mathbf{\Pi}$ replaced by

$$\mathbf{\Pi}_{\text{B.2}} = \Pi'_{\text{B.2}} = \begin{pmatrix} \beta_1\alpha'_1 & 0\cdots0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots0 & \beta_N\alpha'_N \\ \mu_1\alpha'_1 & \cdots & \mu_N\alpha'_N \end{pmatrix}. \tag{50}$$

Minimizing with respect to the β_i's and μ_i's yields the following estimates,

$$\begin{aligned} \begin{pmatrix} \text{vec}(\hat{\gamma}_1) \\ \vdots \\ \text{vec}(\hat{\gamma}_N) \end{pmatrix} &= (\hat{E}'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\hat{E})^{-1}\hat{E}'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\text{vec}(Z'_{-1}\Delta Y) \\ &= (\hat{E}'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes Z'_{-1}Z_{-1})\hat{E}_{\text{SURE}})^{-1}\hat{E}'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes I_T)\text{vec}(Z'_{-1}\Delta Y), \end{aligned} \tag{51}$$

where $\gamma_i = (\beta'_i \ \mu'_i)'$ and $\hat{E}$ is identical to $\hat{E}$ in (31) from section 3.2 with Y_{-1} and $Y_{i,-1}$ replaced by Z_{-1} and $Z_{i,-1} = (Y_{i,-1} \ \ \iota)$ for $i = 1, \ldots, N$. In (51) E_{SURE} equals

$$E_{\text{SURE}} = \left(\ (e_1 \otimes \alpha_1) \otimes \begin{pmatrix} (e_1 \otimes I_k) & \mathbf{0}_{Nk} \\ \mathbf{0}'_k & 1 \end{pmatrix} \ \cdots \ (e_N \otimes \alpha_N) \otimes \begin{pmatrix} (e_N \otimes I_k) & \mathbf{0}_{Nk} \\ \mathbf{0}'_k & 1 \end{pmatrix} \ \right),$$

where $\mathbf{0}_{Nk}$ is a Nk-dimensional vector of zeros and $\mathbf{0}_k$ a k-dimensional vector of zeros.

Next, specification C.2 implies a GMM objective function $G(\Pi_{\text{C.2}},\ \hat{\Omega})$ identical to (49), based on

$$\mathbf{\Pi}_{\text{C.2}} = \Pi'_{\text{C.2}} = \begin{pmatrix} \beta\alpha'_1 & 0\cdots0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots0 & \beta\alpha'_N \\ \mu_1\alpha'_1 & \cdots & \mu_N\alpha'_N \end{pmatrix}. \tag{52}$$

The corresponding GMM-estimates of β and $\mu_1, \ldots, \mu_N$ equals:

$$\begin{aligned} \begin{pmatrix} \text{vec}(\hat{\beta}) \\ \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_N \end{pmatrix} &= (\hat{E}'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\hat{E})^{-1}\hat{E}'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\text{vec}(Z'_{-1}\Delta Y) \\ &= (\hat{E}'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes Z'_{-1}Z_{-1})\hat{E}_{\text{SURE}})^{-1}\hat{E}'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes I_T)\text{vec}(Z'_{-1}\Delta Y), \end{aligned} \tag{53}$$

where

$$\hat{E} = \left(\begin{pmatrix} \hat{\alpha}_1 \otimes Z'_{-1}Y_{1,-1} \\ \vdots \\ \hat{\alpha}_N \otimes Z'_{-1}Y_{N,-1} \end{pmatrix} \begin{pmatrix} (e_1 \otimes \hat{\alpha}_1) \otimes Z'_{-1}\iota & \cdots & (e_N \otimes \hat{\alpha}_N) \otimes Z'_{-1}\iota \end{pmatrix} \right),$$

and

$$\hat{E}_{\text{SURE}} =$$
$$\left(\begin{pmatrix} \hat{\alpha}_1 \otimes \begin{pmatrix} (e_1 \otimes I_k) \\ \mathbf{0}'_k \end{pmatrix} \\ \vdots \\ \hat{\alpha}_N \otimes \begin{pmatrix} (e_N \otimes I_k) \\ \mathbf{0}'_k \end{pmatrix} \end{pmatrix} \begin{pmatrix} (e_1 \otimes \hat{\alpha}_1) \otimes \begin{pmatrix} \mathbf{0}_{Nk} \\ 1 \end{pmatrix} & \cdots & (e_N \otimes \hat{\alpha}_N) \otimes \begin{pmatrix} \mathbf{0}_{Nk} \\ 1 \end{pmatrix} \end{pmatrix} \right).$$

Finally, for specification C.3 GMM objective function $G(\mathbf{\Pi}_{\text{C.3}},\ \hat{\Omega})$ is based on (49) with

$$\mathbf{\Pi}_{\text{C.3}} = \Pi'_{\text{C.3}} = \begin{pmatrix} \beta\alpha'_1 & 0 \cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 \cdots 0 & \beta\alpha'_N \\ \mu\alpha'_1 & \cdots & \mu\alpha'_N \end{pmatrix}. \tag{54}$$

Minimizing $G(\Pi_{\text{C.3}},\ \hat{\Omega})$ with respect to β and μ yields the estimates,

$$\begin{aligned} \text{vec}\begin{pmatrix} \hat{\beta} \\ \hat{\mu} \end{pmatrix} &= (\hat{D}'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\hat{D})^{-1}\hat{D}'(\hat{\Omega} \otimes Z'_{-1}Z_{-1})^{-1}\text{vec}(Z'_{-1}\Delta Y) \\ &= (\hat{D}'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes Z'_{-1}Z_{-1})\hat{D}_{\text{SURE}})^{-1}\hat{D}'_{\text{SURE}}(\hat{\Omega}^{-1} \otimes I_T)\text{vec}(Z'_{-1}\Delta Y), \end{aligned} \tag{55}$$

where $\hat{D}$ is similar to $\hat{D}$ in (35) from section 3.2 with Y_{-1} and $Y_{i,-1}$ replaced by Z_{-1} and $Z_{i,-1}$. Matrix $\hat{D}_{\text{SURE}}$ is based on

$$D_{\text{SURE}} = \begin{pmatrix} \alpha_1 \otimes \begin{pmatrix} (e_1 \otimes I_k) & \mathbf{0}_{Nk} \\ \mathbf{0}'_k & 1 \end{pmatrix} \\ \vdots \\ \alpha_N \otimes \begin{pmatrix} (e_N \otimes I_k) & \mathbf{0}_{Nk} \\ \mathbf{0}'_k & 1 \end{pmatrix} \end{pmatrix}$$

The LR statistics for panel cointegration rank tests based on above mentioned specifications, are in the limit based on the sum of N limiting distributions of the trace statistics for the corresponding deterministic specifications as defined in Johansen (1996, Chapter 6):

Proposition 4.1 *Given the conditions from proposition 3.1, the limiting distribution of LR(*$\Pi_{B.1}|\Pi_{A.1}$*) and LR(*$\Pi_{B.2}|\Pi_{A.1}$*) equals:*

$$LR(\Pi^*_B|\Pi_{A.1}) = T[ln|\hat{\Omega}(\hat{\Pi}^*_B)| - ln|\hat{\Omega}(\hat{\Pi}_{A.1})|] \Rightarrow$$
$$\sum_{i=1}^{N} tr\left(\int dB_{k-r,i}S_i'\left[\int S_iS_i'\right]^{-1}\int S_i dB_{k-r,i}'\right). \quad (56)$$

In (56) $B_{k-r,i}$ *is a* $(k-r)$*-dimensional Brownian motion for individual* i *with an identity covariance matrix and* $\hat{\Omega}(\Pi)$ *is defined in (16), section 3.1.*

B.1*: in (56)* $\mathbf{\Pi}^*_B$ *equals (47) and* S_i *is* $(k-r)$*-dimensional for each individual* i*:*

$$S_i(t) = \begin{pmatrix} B_{k-r-1,i}(t) - \int_0^1 B_{k-r-1,i}(t)dt \\ t - \int_0^1 tdt \end{pmatrix}, \quad (57)$$

where $0 \leq t \leq 1$*.*

B.2*: in (56)* $\mathbf{\Pi}^*_B$ *equals (50) and* S_i *is* $(k-r+1)$*-dimensional for each individual* i*:*

$$S_i(t) = \begin{pmatrix} B_{k-r,i}(t) \\ 1 \end{pmatrix}. \quad (58)$$

Next, we have for LR($\Pi_{C.1}|\Pi_{A.1}$*), LR(*$\Pi_{C.2}|\Pi_{A.1}$*) and LR(*$\Pi_{C.3}|\Pi_{A.1}$*):*

$$LR(\Pi^*_C|\Pi_{A.1}) = T[ln|\hat{\Omega}(\hat{\Pi}^*_C)| - ln|\hat{\Omega}(\hat{\Pi}_{A.1})|] \Rightarrow$$
$$\chi^2(df) + \sum_{i=1}^{N} tr\left(\int dB_{k-r,i}S_i'\left[\int S_iS_i'\right]^{-1}\int S_i dB_{k-r,i}'\right). \quad (59)$$

C.1*: in (59)* $\mathbf{\Pi}^*_C$ *equals (48),* S_i *is identical to (57) and* $\chi^2(df) = \chi^2((N-1)r(k-r))$*.*

C.2*: in (59)* $\mathbf{\Pi}^*_C$ *equals (52),* S_i *is identical to (58) and* $\chi^2(df) = \chi^2((N-1)r(k-r))$*.*

C.3*: in (59)* $\mathbf{\Pi}^*_C$ *equals (54),* S_i *is identical to (58) and, as we have homogeneous restricted constants,* $\chi^2(df) = \chi^2((N-1)r((k-r)+1))$*.*

Proof: *The proofs follow the proofs of propositions 3.2 and 3.3, and are not reproduced.*

4.2 Higher Order Dynamics

When we add individual specific higher order dynamics to the panel VEC models of section 2, our panel VEC models have the following general expression:

$$\Delta Y_t = \Pi^* Y_{t-1} + \Gamma W_t + \varepsilon_t, \quad (60)$$

with the $Nk \times Pk$ matrix

$$\Gamma = \begin{pmatrix} \Gamma_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \Gamma_N \end{pmatrix},$$

where the $k \times p_i k$ matrix Γ_i contains the parameters of the p_i lagged first differences of individual i and $P = \sum_{i=1}^{N} p_i$. We can define W_t in (60) as

$$W_t = \left(\Delta y'_{1,t-1} \cdots \Delta y'_{1,t-p_1} \cdots \Delta y'_{N,t-1} \cdots \Delta y'_{N,t-p_N}\right)',$$

and we define $\Pi^* = \Pi_{\mathrm{A}}$, Π_{B} or Π_{C} as in section 2.

Defining the $T \times Pk$ matrix $W = (W_1 \cdots W_T)'$, $\mathbf{\Gamma} = \Gamma'$ and $U = \Delta Y - Y_{-1}\mathbf{\Pi} - W\mathbf{\Gamma}$, we are able to write the log-likelihood function of (60) as:

$$\begin{aligned} \ell(\Pi^*,\ \Gamma,\ \Omega) &= -\frac{NkT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Omega| - \frac{1}{2}tr\left(\Omega^{-1}U'U\right) \\ &= -\frac{NkT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Omega| - \frac{1}{2}\mathrm{vec}(U)'(\Omega^{-1} \otimes I_T)\mathrm{vec}(U). \end{aligned} \tag{61}$$

One can write in (61) $\mathrm{vec}(U)$ as

$$\begin{aligned} \mathrm{vec}(U) &= \mathrm{vec}(\Delta Y - Y_{-1}\mathbf{\Pi}) - (I_{Nk} \otimes W)\,\mathrm{vec}\,(\mathbf{\Gamma}) \\ &= \mathrm{vec}(\Delta Y - Y_{-1}\mathbf{\Pi}) - (I_{Nk} \otimes W)\,H_{\mathrm{SURE}} \begin{pmatrix} \mathrm{vec}(\Gamma'_1) \\ \vdots \\ \mathrm{vec}(\Gamma'_N) \end{pmatrix}, \end{aligned} \tag{62}$$

with $q_1 = 0$, $q_i = \sum_{j=1}^{i-1} p_j$, $i = 2, \ldots, N$,

$$\begin{aligned} H_{\mathrm{SURE}} &= \begin{pmatrix} (e_1 \otimes I_k) \otimes H_1 & \cdots & (e_N \otimes I_k) \otimes H_N \end{pmatrix}, \\ H_i &= \left(\begin{pmatrix} \mathbf{0}_{q_i} \otimes \mathbf{0}'_{p_i} \\ I_{p_i} \\ \mathbf{0}_{P-p_i-q_i} \otimes \mathbf{0}'_{p_i} \end{pmatrix} \otimes I_k \right);\ i = 1, \ldots, N. \end{aligned}$$

If we concentrate out Γ from (61), we obtain the concentrated log-likelihood of (Π^*, Ω),

$$\ell(\Pi^*\ , \Omega) \propto -\frac{T}{2}\ln|\Omega| - \frac{1}{2}\mathrm{vec}(U)'M_W\mathrm{vec}(U). \tag{63}$$

where,

$$\begin{aligned} M_W = (\Omega^{-1} \otimes I_T) &- (\Omega^{-1} \otimes W)H_{\mathrm{SURE}} \\ &\times \left(H'_{\mathrm{SURE}}(\Omega^{-1} \otimes W'W)H_{\mathrm{SURE}}\right)^{-1} H'_{\mathrm{SURE}}(\Omega^{-1} \otimes W'). \end{aligned}$$

Our GMM objective function then becomes

$$\begin{aligned} G(\mathbf{\Pi},\ \hat{\Omega}) &= \mathrm{vec}(\Delta Y - Y_{-1}\mathbf{\Pi})'M_W(I_{Nk} \otimes Y_{-1}) \\ &\times \left((I_{Nk} \otimes Y_{-1})'M_W(I_{Nk} \otimes Y_{-1})\right)^{-1}(I_{Nk} \otimes Y_{-1})'M_W\mathrm{vec}(\Delta Y - Y_{-1}\mathbf{\Pi}). \end{aligned} \tag{64}$$

The GMM estimates of the cointegrating vectors are now based on (64) and therefore have a slightly different specification. For example, the GMM estimator for the full rank case now equals

$$\begin{pmatrix} \mathrm{vec}(\hat{\Pi}_1') \\ \vdots \\ \mathrm{vec}(\hat{\Pi}_N') \end{pmatrix} = (F_{\mathrm{SURE}}'(I_{Nk} \otimes Y_{-1})' M_W (I_{Nk} \otimes Y_{-1}) F_{\mathrm{SURE}})^{-1} \\ \times F_{\mathrm{SURE}}'(I_{Nk} \otimes Y_{-1})' M_W \mathrm{vec}(\Delta Y), \quad (65)$$

where F_{SURE} is identical to (21) in section 3.2.

Given that

$$(\Omega^{-1} \otimes I_{Nk})\mathrm{vec}(Y_{-1}'\Delta Y) = (I_{Nk} \otimes Y_{-1})'(\Omega^{-1} \otimes I_{Nk})\mathrm{vec}(\Delta Y),$$

one can see that GMM estimator (65) is a straightforward generalization of estimator (22) from section 3.2 to the case of higher order dynamics. The maximum likelihood estimators result from iteratively applying the different GMM estimators and the limiting behavior of the maximum likelihood cointegrating vector estimators therefore remains the same, *i.e.* mixed normal. The limiting distributions of the likelihood ratio statistics based on (63) and (64) are identical to those from propositions 3.2 and 3.3, as higher order dynamics only affect the short-run properties of the model.

5 Cointegration and the Monetary Exchange Rate Model

Our emprical application considers the long-run validity of the monetary exchange rate model. We use a particular version of this model in which the log of the exchange rate at time t (e_t) is related to the differential of the logarithms of the home and foreign money supplies ($m_t - m_t^*$) and the log of relative real income ($y_t - y_t^*$). More specifically,

$$\begin{aligned} e_t &= \mu + (m_t - m_t^*) - \phi(y_t - y_t^*) + \varepsilon_t \\ &= \mu + \tilde{m}_t - \phi\tilde{y}_t + \varepsilon_t, \end{aligned} \quad (66)$$

where $\phi > 0$ and ε_t is a zero mean, stationary deviation at time t with respect to the monetary model. The relationship in (66) has its origin in the work of Mussa (1976).

As the variables in (66) are known to be I(1) processes, we can analyze the monetary model in a VEC framework. Within an individual VEC model validity of the monetary exchange rate model implies for country i:

$$\Delta x_{it} = \begin{pmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \alpha_{i3} \end{pmatrix} \begin{pmatrix} \beta^{r'} & -\mu_i \end{pmatrix} z_{i,t-1} + \varepsilon_{it}, \quad (67)$$

where $\Delta x_t = (\Delta e_{it} \;\; \Delta\tilde{m}_{it} \;\; \Delta\tilde{y}_{it})'$, $z_{i,t-1} = (e_{i,t-1} \;\; \tilde{m}_{i,t-1} \;\; \tilde{y}_{i,t-1} \;\; 1)'$, $\beta^r = (1 \;\; -1 \;\; \phi_i)'$ and ε_{it} is a vector of white noise disturbances. Hence, the monetary model in (66) implies for a VEC model of *each* country a reduced rank value equal to 1, equality of the parameter values of e_{it} and $\tilde{m}_{it}$ in absolute terms with opposite signs within the cointegrating vector, and a positive value of the income elasticity ϕ_i. The last restriction fulfills the condition $-\phi < 0$ in the equilibrium relationship (66).

We analyze in our application US dollar rates and the corresponding monetary fundamentals relative to the US for France, Germany and the United Kingdom (UK), which are the three major European economies. The data are quarterly and start in the first quarter of 1973, with the abolishment of the Bretton Woods system, and ends in the last quarter of 1994. As the money supply measure we use seasonally unadjusted M1 aggregates and for real income we use Gross Domestic Product (GDP), or Gross National Product (GNP) in the case of Germany. The data are retrieved from the *International Financial Statistics* with the exception of Germany (GNP from *Main Economic Indicators*) and the UK (M1 from the Dutch Central Bank). See also Groen (1999, Appendix A) for a more detailed description of the data.

We first test for cointegration for each country separately using Johansen's maximum likelihood approach. To guarantee white noise residuals in the individual VEC models we allow for higher order dynamics and the lag order is selected based on information criteria and white noise residual tests. Seasonal patterns are corrected for through the usage of three zero mean seasonal dummies in the VEC models. Table 1 contains the results of the cointegration tests for the analyzed data. One can infer from these results that there is no evidence for the appropriateness of the monetary model for long-run developments in US dollar nominal exchange rates of France, Germany and the UK. In non of the cases can we reject at a 5% significance level the null hypothesis of a reduced rank value equal to 0, *i.e.* the null hypothesis of no cointegration. These results corroborate the time series-based cointegration test results of amongst others Sarantis (1994) and Groen (1999).

As discussed before, the monetary exchange rate model implies a common structure for all the countries in our data-set. Next to that, the US dollar exchange rates of European countries have high cross-country correlations. Based on these considerations Groen (1999) applied the panel Engle-Granger approach on the monetary model with homogeneous long-run parameters for $\tilde{m}_{it}$ and $\tilde{y}_{it}$, and he indeed finds evidence for cointegration within panels of at least nine countries. Hence, it seems worthwhile to analyze the variables of France, Germany and the UK within a panel of three VEC models.

Using the different specifications of the cointegrating vectors and deterministic components discussed in sections 2 and 4.1, we test the following hypotheses within our three country panel VEC model:

- **B.2**$(r)|$**A.1** tests $\mathbf{\Pi}_{\mathrm{B.2}}$ (50) under common cointegration rank r versus $\mathbf{\Pi}_{\mathrm{A.1}}$ (45).
- **C.2**$(r)|$**A.1** tests $\mathbf{\Pi}_{\mathrm{C.2}}$ (52) under common cointegration rank r versus $\mathbf{\Pi}_{\mathrm{A.1}}$ (45).
- **C.2H**$(r)|$**A.1** tests $\mathbf{\Pi}_{\mathrm{C.2H}}$ under common cointegration rank r versus $\mathbf{\Pi}_{\mathrm{A.1}}$ (45),

Table 1: Cointegration tests for the individual countries, 1973:1-1994:4

lags	rank 0	rank 1	rank 2
		France	
3	32.16	15.49	5.55
		Germany	
3	16.82	9.12	3.83
		United Kingdom	
3	20.85	9.59	2.14
		5% Critical Values[a]	
–	34.91	19.96	9.24
		1% Critical Values[a]	
–	41.07	24.60	12.97

[a] See Osterwald-Lenum (1992).

where $\mathbf{\Pi}_{\text{C.2H}}$ equals (52) with the parameter of $\tilde{y}_{it}$ in the cointegrating vector assumed to be heterogeneous across the countries.

The above mentioned hypotheses are tested with the likelihood ratio statistics (LR) that are based on the GMM estimators iterated over *all* the parameters, including the disturbance covariance matrix, resulting in maximum likelihood estimates. Appropriate limiting distributions are summarized in proposition 4.1. The corresponding critical values are computed using simulations based on 1,000 time series observations and 100,000 iterations, where the procedures in Johansen (1996) are used for the individual $(k-r)$-dimensional Brownian motions. Note that the value of the degrees of freedom of the χ^2 part of the limiting distribution of $\mathbf{C.2H}(r)|\mathbf{A.1}$ equals $(N-1)r(k-r-1)$ instead of $(N-1)r(k-r)$ for $\mathbf{C.2}(r)|\mathbf{A.1}$.

The corresponding test results for our data-set can be found in table 2. Within our panel VEC models we make use of the same lag order for the lagged first differences as in the individual cointegration analysis and again we use three zero mean seasonal dummies. The results in table 2 indicate that we can reject the hypothesis of no cointegration across all countries. We also cannot reject the hypothesis of a common cointegration rank value equal to 1 based on both fully and partial ($=\mathbf{\Pi}_{\text{C.2H}}$) heterogeneous cointegrating vectors. The LR statistic for specification **C.2** indicate that we cannot reject at a 5% significance level the presence of one common cointegration vector based on homogeneous long-run parameters of e_{it}, $\tilde{m}_{it}$ and $\tilde{y}_{it}$, but the statistic rejects the validity of specification **C.2** at the 10% critical value equal to 50.69. As we use the asymptotic distribution and not the

Table 2: Cointegration rank tests in the panel VEC models, 1973:1-1994:4

	LR	95%[a]	99%[a]
B.2(0)|A.1	95.64*	90.64	101.20
B.2(1)|A.1	28.55	49.61	56.06
B.2(2)|A.1	3.84	20.74	25.16
C.2(1)|A.1	51.30	54.27	61.84
C.2(2)|A.1	17.21	26.59	31.56
C.2H(1)|A.1	43.64	52.31	59.32
C.2H(2)|A.1	3.25	20.74	25.16

[a] "95%" ("99%") are the 95% (99%) quantiles of the appropriate limiting distribution and *(**) indicates a rejection of the null hypothesis at these quantiles.

actual distribution of the test statistic we have a weak rejection of **C.2**. Thus we cannot make a clear decision whether to accept or reject **C.2** as a valid specification.

The normalized cointegating vectors which corresponds with the specifications that we could not reject in table 2 are reported in table 3. The estimated cointegrating vector parameters of log relative real income $\tilde{y}_{it}$ have the proper signs as in (67), both when assumed to be heterogeneous and homogeneous. Specification (67) also implies for all countries that the long-run parameters of e_{it} and $\tilde{m}_{it}$ must have equal absolute values with opposite signs. Therefore, $(\hat{\beta}_e + \hat{\beta}_{\tilde{m}}) = 0$ should be a valid restriction on the estimated cointegrating vectors $\hat{\beta} = (\hat{\beta}_e \;\; \hat{\beta}_{\tilde{m}} \;\; \hat{\beta}_{\tilde{y}} \;\; \hat{\beta}_{\mu,i})'$ or $\hat{\beta} = (\hat{\beta}_e \;\; \hat{\beta}_{\tilde{m}} \;\; \hat{\beta}_{\tilde{y},i} \;\; \hat{\beta}_{\mu,i})'$ for $i = 1, 2, 3$. We test the restriction $(\hat{\beta}_e + \hat{\beta}_{\tilde{m}}) = 0$ conditional on homogeneous long-run parameters of e_{it} and $\tilde{m}_{it}$ and a given cointegration rank. Hence, the corresponding LR statistics have a limiting distribution equal to a $\chi^2(1)$ distribution. The results in table 3 indicate that in none of the cases we are able to reject the null hypothesis $(\hat{\beta}_1 + \hat{\beta}_2) = 0$, rending support to the cointegrating vector as implied by the monetary exchange rate model in (67).

6 Conclusions

In this paper we construct a framework for cointegration analysis in panels with a fixed number of vector error correction models. As analytical expressions of maximum likelihood estimators based on canonical correlations cannot be constructed within our panel vector error correction model, we use maximum likelihood estimates that result from an iterative estimation procedure based on GMM expressions of the parameters.

The GMM estimators result from an objective function that is embedded in the log-likelihood function and these estimators can be interpreted as SURE-type estimators.

Table 3: Normalized cointegrating vectors within the panel VEC models

	$\beta^{a}_{C.2}$	$\beta^{a}_{C.2H}$
e_{it}	1	1
$\tilde{m}_{it}$	-1.23	-0.90
$\tilde{y}_{it}$	5.58	–
$\tilde{y}_{FR,t}$	–	5.15
$\tilde{y}_{GER,t}$	–	3.26
$\tilde{y}_{UK,t}$	–	11.16
μ_{FR}	-2.10	-2.19
μ_{GER}	2.69	1.13
μ_{UK}	9.91	23.78
$(\hat{\beta}_e + \hat{\beta}_{\tilde{m}}) = 0^{b}$	1.40	0.22
	(0.28)	(0.64)

[a] Specifications C.2 and C.2H are defined in the text.

[b] LR statistic testing that e_{it} and $\tilde{m}_{it}$ have opposite parameter values with the corresponding $\chi^2(1)$ p-values in parentheses.

Iteratively applying the GMM estimators for the different parameters yields maximum likelihood estimates after convergence of the estimators and the GMM objective function. We use these maximum likelihood estimates to construct likelihood ratio statistics to test for a common cointegration rank value across the individual vector error correction models within our panel.

Our likelihood ratio tests for a common cointegration rank have a limiting distribution equal to a summation of limiting distributions of an appropriate number of Johansen (1996) trace statistics. When we assume that the cointegrating vectors are (partly) homogeneous across the individuals, we see that the limiting distributions of the rank test statistics are composed of a summation of Brownian motion functionals and a χ^2 random variable. We also formulate cointegrating vector estimators and rank test statistics for different specifications of the deterministic components. These are derived in similar way as before.

To show the importance of exploiting common structures across different individual vector error correction models, we analyze the appropriateness of the monetary exchange rate model for US dollar exchange rates. Within our panel-structure approach the test results indicates strong support for the validity of the monetary exchange rate model for long-run developments in the US dollar exchange rates of France, Germany and the UK. This result is indicative for the importance of using the additional information in the cross-section dimension of economic time-series.

For future research there are many applications where our method can make a difference. Some examples are, purchasing power parity as in *e.g.* Pedroni (1995, 1996) and long run import demand relationships as in Kleibergen *et al.* (1999). Economic applications like the above mentioned offer interesting areas for future research. One should realize that even when common structures do not exist across individuals, it is often more efficient to estimate a panel VEC model based on heterogeneous cointegrating vectors when we have cross-section correlations. Individual vector error correction models are only asymptotically efficient in the absence of cross-sectional correlation. When cross-sectional correlation exists, system methods, like the panel VEC model with heterogeneous cointegrating vectors, are asymptotically more efficient.

Appendix

A Representation Theorem

We can directly apply the Granger-Johansen representation theorem, see Johansen (1991), to the system of the cointegrated VEC models,

$$\Delta Y_t = \begin{pmatrix} \alpha_1'\beta_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \alpha_N'\beta_N \end{pmatrix} Y_{t-1} + \varepsilon_t,$$

and obtain its stochastic trend specification,

$$Y_t = \begin{pmatrix} \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}' & 0\ldots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\ldots 0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}' \end{pmatrix} \sum_{s=0}^{t-1} \varepsilon_t + Z_t,$$

where Z_t is a Nk dimensional stationary time series and $\beta_{i\perp}'\beta_i \equiv 0$, $\alpha_{i\perp}'\alpha_i \equiv 0$, $i = 1, \ldots, N$. The limiting behavior of Y_t is therefore such that

$$\frac{1}{\sqrt{T}} Y_t \Rightarrow \begin{pmatrix} \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0\ldots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\ldots 0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix} \Lambda^{\frac{1}{2}} B_{N(k-r)},$$

where $B_{N(k-r)}$ is a $N(k-r)$ dimensional Brownian motion with an identity covariance matrix and

$$\begin{aligned} \Lambda &= \begin{pmatrix} \alpha_{1\perp}' & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \alpha_{N\perp}' \end{pmatrix} \Omega \begin{pmatrix} \alpha_{1\perp} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \alpha_{N\perp} \end{pmatrix} \\ &= \begin{pmatrix} (e_1 \otimes \alpha_{1\perp}) & \cdots & (e_N \otimes \alpha_{N\perp}) \end{pmatrix}' \Omega \begin{pmatrix} (e_1 \otimes \alpha_{1\perp}) & \cdots & (e_N \otimes \alpha_{N\perp}) \end{pmatrix}, \end{aligned}$$

where e_i is the i-th N dimensional unity vector.

B Proof of Proposition 3.1

The maximum likelihood estimate of the cointegrating vectors $\beta_1, \ldots, \beta_N$ results from iteratively applying the cointegrating vector estimator (31) based on consistent estimates of $\alpha_1, \ldots, \alpha_N$ and the disturbance covariance matrix Ω. The error introduced by using these consistent estimators instead of the true unknown value is, however, such that the highest order in the number of observations of the limiting expression of the cointegrating vector estimator is not affected by it. Hence, to construct the limiting distribution of the cointegrating vector estimator, we can treat the consistent estimators as if they are equal to the true value of the parameter.

The GMM cointegrating vector estimator which equals, after convergence, the maximum likelihood estimates of $\beta_1, \ldots, \beta_N$ reads,

$$\begin{pmatrix} \text{vec}(\hat{\beta}_1) \\ \vdots \\ \text{vec}(\hat{\beta}_N) \end{pmatrix} = (E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})E)^{-1}E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\text{vec}(Y_{-1}'\Delta Y),$$

where $Y_{-1} = (Y_{1,-1} \cdots Y_{N,-1})$, $Y_{i,-1}$ is $T \times k$, $i = 1, \ldots, N$, and

$$\begin{aligned} E &= \begin{pmatrix} (e_1 \otimes \hat{\alpha}_1) \otimes (Y_{-1}'Y_{1,-1}) & \cdots & (e_N \otimes \hat{\alpha}_N) \otimes (Y_{-1}'Y_{N,-1}) \end{pmatrix} \\ &= \left((I_N \otimes I_k) \otimes Y_{-1}'Y_{-1}\right) \begin{pmatrix} (e_1 \otimes \tilde{\alpha}_1) \otimes (e_1 \otimes I_k) & \cdots & (e_N \otimes \tilde{\alpha}_N) \otimes (e_N \otimes I_k) \end{pmatrix}, \end{aligned}$$

with $\hat{\Omega}$ and $\hat{\alpha}_1, \ldots, \hat{\alpha}_N$ as consistent estimates of Ω and $\alpha_1, \ldots, \alpha_N$. We substitute that $\Delta Y = Y_{-1}\text{diag}(\beta_i\alpha_i') + \varepsilon$, where $\text{diag}(A_i)$ is a diagonal matrix with A_i on the diagonal, such that

$$\begin{aligned} \begin{pmatrix} \text{vec}(\hat{\beta}_1) \\ \vdots \\ \text{vec}(\hat{\beta}_N) \end{pmatrix} &= (E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})E)^{-1}E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\text{vec}(Y_{-1}'\Delta Y) \\ &= (E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})E)^{-1}E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1}) \left[E \begin{pmatrix} \text{vec}(\beta_1) \\ \vdots \\ \text{vec}(\beta_N) \end{pmatrix} + \text{vec}(Y_{-1}'\varepsilon) \right] \\ &= \begin{pmatrix} \text{vec}(\beta_1) \\ \vdots \\ \text{vec}(\beta_N) \end{pmatrix} + (E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})E)^{-1}E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\text{vec}(Y_{-1}'\varepsilon). \end{aligned}$$

The limiting behavior of the cointegrating vector estimator now results from the limiting behavior of the different elements of the last part of the above expression. We therefore construct the limiting behavior of each of these different elements.

$$\begin{aligned} &\frac{1}{T^2}E \Rightarrow \\ &\left((I_N \otimes I_k) \otimes \begin{pmatrix} \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0\ldots0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\ldots0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix}\right) \\ &\left((I_N \otimes I_k) \otimes \Lambda^{\frac{1}{2}} \left(\int B_{N(k-r)}B_{N(k-r)}'\right) \Lambda^{\frac{1}{2}\prime}\right) \\ &\left((I_N \otimes I_k) \otimes \begin{pmatrix} \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0\ldots0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\ldots0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix}'\right) \\ &\begin{pmatrix} (e_1 \otimes \alpha_1) \otimes (e_1 \otimes I_k) & \cdots & (e_N \otimes \alpha_N) \otimes (e_N \otimes I_k) \end{pmatrix}, \end{aligned}$$

and

$$\frac{1}{T^2}Y'_{-1}Y_{-1} \Rightarrow$$

$$\begin{pmatrix} \beta_{1\perp}(\alpha'_{1\perp}\beta_{1\perp})^{-1} & 0\dots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\dots 0 & \beta_{N\perp}(\alpha'_{N\perp}\beta_{N\perp})^{-1} \end{pmatrix} \Lambda^{\frac{1}{2}} \left(\int B_{N(k-r)}B'_{N(k-r)} \right) \Lambda^{\frac{1}{2}\prime}$$
$$\begin{pmatrix} \beta_{1\perp}(\alpha'_{1\perp}\beta_{1\perp})^{-1} & 0\dots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\dots 0 & \beta_{N\perp}(\alpha'_{N\perp}\beta_{N\perp})^{-1} \end{pmatrix}',$$

such that

$$T^2(Y'_{-1}Y_{-1})^{-1} \Rightarrow$$

$$\begin{pmatrix} \beta_{1\perp}(\beta'_{1\perp}\beta_{1\perp})^{-1}(\alpha'_{1\perp}\beta_{1\perp}) & 0\dots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \beta_{N\perp}(\beta'_{N\perp}\beta_{N\perp})^{-1}(\alpha'_{N\perp}\beta_{N\perp}) \end{pmatrix}$$
$$\Lambda^{-\frac{1}{2}\prime} \left(\int B_{N(k-r)}B'_{N(k-r)} \right)^{-1} \Lambda^{-\frac{1}{2}}$$
$$\begin{pmatrix} \beta_{1\perp}(\beta'_{1\perp}\beta_{1\perp})^{-1}(\alpha'_{1\perp}\beta_{1\perp}) & 0\dots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\dots 0 & \beta_{N\perp}(\beta'_{N\perp}\beta_{N\perp})^{-1}(\alpha'_{N\perp}\beta_{N\perp}) \end{pmatrix}'.$$

As a consequence of this,

$$\frac{1}{T^2}E'(\hat{\Omega}^{-1} \otimes (Y'_{-1}Y_{-1})^{-1})E \Rightarrow$$

$$\begin{pmatrix} (e_1 \otimes \alpha_1) \otimes (e_1 \otimes I_k) & \cdots & (e_N \otimes \alpha_N) \otimes (e_N \otimes I_k) \end{pmatrix}'$$
$$\left((I_N \otimes I_k) \otimes \begin{pmatrix} \beta_{1\perp}(\alpha'_{1\perp}\beta_{1\perp})^{-1} & 0\dots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\dots 0 & \beta_{N\perp}(\alpha'_{N\perp}\beta_{N\perp})^{-1} \end{pmatrix} \right)$$
$$\left(\Omega^{-1} \otimes \Lambda^{\frac{1}{2}} \left(\int B_{N(k-r)}B'_{N(k-r)} \right) \Lambda^{\frac{1}{2}\prime} \right)$$
$$\left((I_N \otimes I_k) \otimes \begin{pmatrix} \beta_{1\perp}(\alpha'_{1\perp}\beta_{1\perp})^{-1} & 0\dots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\dots 0 & \beta_{N\perp}(\alpha'_{N\perp}\beta_{N\perp})^{-1} \end{pmatrix}' \right)$$
$$\begin{pmatrix} (e_1 \otimes \alpha_1) \otimes (e_1 \otimes I_k) & \cdots & (e_N \otimes \alpha_N) \otimes (e_N \otimes I_k) \end{pmatrix} =$$

$$\begin{pmatrix} I_r\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix} \begin{pmatrix} (e_1\otimes\alpha_1)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes\alpha_N)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}$$

$$\left(\Omega^{-1}\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)$$

$$\begin{pmatrix} (e_1\otimes\alpha_1)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes\alpha_N)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}' \begin{pmatrix} I_r\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix}' =$$

$$\begin{pmatrix} I_r\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix} \begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}$$

$$\left(\Psi\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)$$

$$\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}' \begin{pmatrix} I_r\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix}',$$

where $\Psi = \begin{pmatrix} (e_1\otimes\alpha_1) & \cdots & (e_N\otimes\alpha_N) \end{pmatrix}'\Omega^{-1}\begin{pmatrix} (e_1\otimes\alpha_1) & \cdots & (e_N\otimes\alpha_N) \end{pmatrix}$. The limiting behavior of the inverse of the latter expression is therefore characterized by

$$T^2\left(E'(\hat\Omega^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})E\right)^{-1}\Rightarrow$$

$$\begin{pmatrix} I_r\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}$$

$$\left[\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}\left(\Psi\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)\right.$$

$$\left.\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\right]^{-1}$$

$$\begin{pmatrix} I_r\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}'.$$

The limiting behavior of $\mathrm{vec}(Y_{-1}'\varepsilon)$ is characterized by

$$\frac{1}{T}\mathrm{vec}(Y_{-1}'\varepsilon)\Rightarrow$$

$$\left(\Omega^{\frac{1}{2}}\otimes\begin{pmatrix} \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}' & 0\ldots0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\ldots0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}' \end{pmatrix}\Omega^{\frac{1}{2}}\right) vec\left(\int B_{Nk}dB_{Nk}'\right).$$

such that the limiting behavior of $E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\text{vec}(Y_{-1}'\varepsilon)$ is characterized by

$$\frac{1}{T}E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\text{vec}(Y_{-1}'\varepsilon) \Rightarrow$$

$$\begin{pmatrix} (e_1 \otimes \alpha_1)' \otimes \left(e_1' \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\right) \\ \vdots \\ (e_N \otimes \alpha_N)' \otimes \left(e_N' \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\right) \end{pmatrix}$$

$$\left(\Omega^{-\frac{1}{2}} \otimes \begin{pmatrix} \alpha_{1\perp}' & 0\ldots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\ldots 0 & \alpha_{N\perp}' \end{pmatrix} \Omega^{\frac{1}{2}}\right) vec\left(\int B_{Nk}dB_{Nk}'\right) =$$

$$\begin{pmatrix} (e_1 \otimes \alpha_1)'\Omega^{-\frac{1}{2}} \otimes \left(e_1' \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\right)\Omega^{\frac{1}{2}} \\ \vdots \\ (e_N \otimes \alpha_N)'\Omega^{-\frac{1}{2}} \otimes \left(e_N' \otimes \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\right)\Omega^{\frac{1}{2}} \end{pmatrix} vec\left(\int B_{Nk}dB_{Nk}'\right) \Rightarrow$$

$$\begin{pmatrix} vec\left(\left(e_1' \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\right)\Omega^{\frac{1}{2}} \int B_{Nk}dB_{Nk}'\Omega^{-\frac{1}{2}\prime}(e_1 \otimes \alpha_1)\right) \\ \vdots \\ vec\left(\left(e_N' \otimes \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\right)\Omega^{\frac{1}{2}} \int B_{Nk}dB_{Nk}'\Omega^{-\frac{1}{2}\prime}(e_N \otimes \alpha_N)\right) \end{pmatrix}.$$

As $\left(e_i \otimes \beta_{i\perp}(\alpha_{i\perp}'\beta_{i\perp})^{-1}\alpha_{i\perp}\right)'\Omega^{\frac{1}{2}}\Omega^{-\frac{1}{2}}(e_i \otimes \alpha_i) = 0$, the Brownian motions involved in the above expression are stochastically independent, see Phillips (1991), and the limiting distribution of the cointegrating vector estimator is therefore mixed normal. Furthermore,

$$\frac{1}{T}E'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\text{vec}(Y_{-1}'\varepsilon) \Rightarrow$$

$$\begin{pmatrix} I_r \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r \otimes \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix} vec\left(\Lambda^{\frac{1}{2}} \int B_{N(k-r)}dB_{Nr}'\Psi^{\frac{1}{2}\prime}\right),$$

such that the limiting behavior of the cointegrating vector estimator is characterized by

$$
T\begin{pmatrix} \mathrm{vec}(\hat{\beta}_1-\beta_1) \\ \vdots \\ \mathrm{vec}(\hat{\beta}_N-\beta_N) \end{pmatrix} \Rightarrow
$$

$$
\begin{pmatrix} I_r\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}
$$

$$
\left[\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}\left(\Psi\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)\right.
$$

$$
\left.\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\right]^{-1}
$$

$$
\begin{pmatrix} I_r\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}'
$$

$$
\begin{pmatrix} I_r\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix} vec\left(\Lambda^{\frac{1}{2}}\int B_{N(k-r)}dB_{Nr}'\Psi^{\frac{1}{2}\prime}\right) =
$$

$$
\begin{pmatrix} I_r\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}
$$

$$
\left[\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}\left(\Psi\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\right]^{-1}
$$

$$
\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1\otimes I_{k-r})' \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N\otimes I_{k-r})' \end{pmatrix} vec\left(\Lambda^{\frac{1}{2}}\int B_{N(k-r)}dB_{Nr}'\Psi^{\frac{1}{2}\prime}\right).
$$

Although we can proof the asymptotic normality of the cointegrating vector estimator, as $B_{N(k-r)}$ and B_{Nr} in the above expression are stochastically independent, we cannot construct a convenient expression of the asymptotic covariance matrix. This is a consequence of the fact that

$$
\begin{pmatrix} (e_1\otimes I_r)\otimes(e_1\otimes I_{k-r}) & \cdots & (e_N\otimes I_r)\otimes(e_N\otimes I_{k-r}) \end{pmatrix}
$$

is not invertible, as it is not a square matrix, which is caused by the diagonal structure of the long run multiplier.

As we normalized β_i as $\beta_i=(I_r\ -\beta_{2i}')'$, the limiting distribution in the above expression is degenerate.

The limiting distribution of the $\hat{\beta}_{2i}$, that is not degenerate is characterized by

$$T \begin{pmatrix} \mathrm{vec}(\hat{\beta}_{21}-\beta_{21}) \\ \vdots \\ \mathrm{vec}(\hat{\beta}_{2N}-\beta_{2N}) \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} I_r \otimes (\beta'_{1\perp}\beta_{1\perp})^{-1}\alpha'_{1\perp}\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_r \otimes (\beta'_{N\perp}\beta_{N\perp})^{-1}\alpha'_{N\perp}\beta_{N\perp} \end{pmatrix}$$

$$\left[\begin{pmatrix} (e_1\otimes I_r)'\otimes(e'_1\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e'_N\otimes I_{k-r}) \end{pmatrix} \left(\Psi\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B'_{N(k-r)}\right)\Lambda^{\frac{1}{2}\prime}\right) \begin{pmatrix} (e_1\otimes I_r)'\otimes(e'_1\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_r)'\otimes(e'_N\otimes I_{k-r}) \end{pmatrix}'\right]^{-1}$$

$$\begin{pmatrix} (e_1\otimes I_r)'\otimes(e_1\otimes I_{k-r})' \\ \vdots \\ (e_N\otimes I_r)'\otimes(e_N\otimes I_{k-r})' \end{pmatrix} \mathrm{vec}\left(\Lambda^{\frac{1}{2}}\int B_{N(k-r)}dB'_{Nr}\Psi^{\frac{1}{2}\prime}\right),$$

as $\beta_{i\perp} = (\beta_{2i}\ I_{k-r})'$.

C Proof of Proposition 3.2

Substituting the true disturbance covariance matrix Ω in log-likelihood function (18) from section 3.1 for both the Π_{B} and Π_{A} specifications makes it possible to rewrite the LR test statistic as,

$$\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) = 2\left[\ell(\hat{\Pi}_{\mathrm{A}},\ \Omega) - \ell(\hat{\Pi}_{\mathrm{B}},\ \Omega)\right] = G(\mathbf{\Pi}_{\mathrm{B}},\Omega) - G(\mathbf{\Pi}_{\mathrm{A}}\ ,\Omega),$$

where

$$G(\mathbf{\Pi}_{\mathrm{B}},\ \Omega) =$$
$$\mathrm{vec}\left(Y'_{-1}\Delta Y - E\begin{pmatrix}\mathrm{vec}(\hat{\beta}_1)\\ \vdots \\ \mathrm{vec}(\hat{\beta}_N)\end{pmatrix}\right)'\left(\Omega^{-1}\otimes(Y'_{-1}Y_{-1})^{-1}\right)\mathrm{vec}\left(Y'_{-1}\Delta Y - E\begin{pmatrix}\mathrm{vec}(\hat{\beta}_1)\\ \vdots \\ \mathrm{vec}(\hat{\beta}_N)\end{pmatrix}\right),$$

and

$$G(\mathbf{\Pi}_{\mathrm{A}},\ \Omega) =$$
$$\mathrm{vec}\left(Y'_{-1}\Delta Y - F\begin{pmatrix}\mathrm{vec}(\hat{\Pi}_1)\\ \vdots \\ \mathrm{vec}(\hat{\Pi}_N)\end{pmatrix}\right)'\left(\Omega^{-1}\otimes(Y'_{-1}Y_{-1})^{-1}\right)\mathrm{vec}\left(Y'_{-1}\Delta Y - F\begin{pmatrix}\mathrm{vec}(\hat{\Pi}_1)\\ \vdots \\ \mathrm{vec}(\hat{\Pi}_N)\end{pmatrix}\right),$$

with

$$\begin{aligned} F &= \begin{pmatrix} (e_1\otimes I_k)\otimes(Y'_{-1}Y_{1,-1}) & \cdots & (e_N\otimes I_k)\otimes(Y'_{-1}Y_{N,-1}) \end{pmatrix} \\ &= \left((I_N\otimes I_k)\otimes Y'_{-1}Y_{-1}\right)\begin{pmatrix} (e_1\otimes I_k)\otimes(e_1\otimes I_k) & \cdots & (e_N\otimes I_k)\otimes(e_N\otimes I_k) \end{pmatrix}. \end{aligned}$$

Under a true common cointegration rank value r $\hat{\Omega}(\Pi_{\mathrm{A}})$, as defined in (25) in section 3.2, is a consistent estimate of Ω. Based on the mapping theorem of Billingsley (1986, Theorem 25.7, Corollary

2) and the line of reasoning at the beginning of Appendix B, we have that $\ell(\hat{\Pi}_{\mathrm{A}},\ \hat{\Omega}(\Pi_{\mathrm{A}})) \Rightarrow \ell(\hat{\Pi}_{\mathrm{A}},\ \Omega)$ and $\ell(\hat{\Pi}_{\mathrm{B}},\ \hat{\Omega}(\Pi_{\mathrm{A}})) \Rightarrow \ell(\hat{\Pi}_{\mathrm{B}},\ \Omega)$ conditional on $\hat{\Omega}(\Pi_{\mathrm{A}}) \Rightarrow \Omega$. Therefore, we can approximate $\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}})$ as

$$LR(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) \backsimeq G(\mathbf{\Pi}_{\mathrm{B}},\hat{\Omega}) - G(\mathbf{\Pi}_{\mathrm{A}}\ ,\hat{\Omega}),$$

with $\hat{\Omega} = \hat{\Omega}(\Pi_{\mathrm{A}})$.

The limiting behavior of $\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}})$ results from constructing the limiting behavior of the different elements of $G(\mathbf{\Pi}_{\mathrm{B}},\hat{\Omega}) - G(\mathbf{\Pi}_{\mathrm{A}}\ ,\hat{\Omega})$. Consequently, we first analyze these different elements. We have:

$$\begin{aligned}\mathrm{vec}\left(Y_{-1}'\Delta Y - E\begin{pmatrix}\mathrm{vec}(\hat{\beta}_1)\\ \vdots \\ \mathrm{vec}(\hat{\beta}_N)\end{pmatrix}\right) = \\ \left(I_{Nk} - E(E'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})E)^{-1}E'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})\right)\mathrm{vec}\left(Y_{-1}'\Delta Y\right),\end{aligned}$$

where we have substituted GMM estimator (31) based on $\hat{\Omega} = \hat{\Omega}(\Pi_{\mathrm{A}})$. The usage of consistent estimators for $\alpha_1, \ldots, \alpha_N$ and Ω in the GMM objective function for the Π_{B} specification is asymptotically for reasons explained in the beginning of Appendix B and the mapping theorem of Billingsley (1986, Theorem 25.7, Corollary 2). The aforementioned leads to

$$\begin{aligned}G(\mathbf{\Pi}_{\mathrm{B}},\ \hat{\Omega}) &= \mathrm{vec}\left(Y_{-1}'\Delta Y\right)'\Big((\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1}) - (\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})E \\ &\qquad (E'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})E)^{-1}E'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})\Big)\mathrm{vec}\left(Y_{-1}'\Delta Y\right) \\ &= \mathrm{vec}\left(Y_{-1}'\Delta Y\right)'E_{\perp}\left(E_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))E_{\perp}\right)^{-1}E_{\perp}'\mathrm{vec}\left(Y_{-1}'\Delta Y\right),\end{aligned}$$

and

$$\begin{aligned}G(\mathbf{\Pi}_{\mathrm{A}},\ \hat{\Omega}) &= \mathrm{vec}\left(Y_{-1}'\Delta Y\right)'\Big((\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1}) - (\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F \\ &\qquad (F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F)^{-1}F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})\Big)\mathrm{vec}\left(Y_{-1}'\Delta Y\right) \\ &= \mathrm{vec}\left(Y_{-1}'\Delta Y\right)'F_{\perp}\left(F_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))F_{\perp}\right)^{-1}F_{\perp}'\mathrm{vec}\left(Y_{-1}'\Delta Y\right),\end{aligned}$$

where $E_{\perp}'E \equiv 0$, $F_{\perp}'F \equiv 0$. As $E = FH$, where

$$\begin{aligned}H &= \begin{pmatrix}(1\otimes\alpha_1)\otimes I_k & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & (1\otimes\alpha_N)\otimes I_k\end{pmatrix} : Nk^2\times Nkr, \\ H_{\perp} &= \begin{pmatrix}(1\otimes\alpha_{1\perp})\otimes I_k & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & (1\otimes\alpha_{N\perp})\otimes I_k\end{pmatrix} : Nk^2\times Nk(k-r),\end{aligned}$$

a convenient specification of $E_{\perp}$ is,

$$E_{\perp} = \begin{pmatrix} F_{\perp} & (\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp}\end{pmatrix}.$$

$E_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))E_{\perp}$ can then be specified as,

$$E_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))E_{\perp} = \begin{pmatrix}F_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))F_{\perp} & 0 \\ 0 & H_{\perp}'(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp}\end{pmatrix}$$

and thus

$$E_{\perp}\left(E_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))E_{\perp}\right)^{-1}E_{\perp}' = F_{\perp}\left(F_{\perp}'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))F_{\perp}\right)^{-1}F_{\perp}' +$$
$$(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp}\left(H_{\perp}'(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp}\right)^{-1}H_{\perp}'$$
$$(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1},$$

such that

$$\begin{aligned}
\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) &\backsimeq G(\mathbf{\Pi}_{\mathrm{B}},\ \hat{\Omega}) - G(\mathbf{\Pi}_{\mathrm{A}},\ \hat{\Omega}) \\
&\backsimeq \mathrm{vec}(Y_{-1}'\Delta Y)'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp} \\
&\qquad \left(H_{\perp}'(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp}\right)^{-1}H_{\perp}' \\
&\qquad\qquad (F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}\mathrm{vec}(Y_{-1}'\Delta Y) \\
&\backsimeq \mathrm{vec}(Y_{-1}'\varepsilon)'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp} \\
&\qquad \left(H_{\perp}'(F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}H_{\perp}\right)^{-1}H_{\perp}' \\
&\qquad\qquad (F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}F)^{-1}F'(\hat{\Omega}\otimes(Y_{-1}'Y_{-1}))^{-1}\mathrm{vec}(Y_{-1}'\varepsilon),
\end{aligned}$$

because $\mathrm{vec}(Y_{-1}'\Delta Y) = FH\mathrm{vec}\begin{pmatrix}\beta_1\\ \vdots \\ \beta_N\end{pmatrix} + \mathrm{vec}(Y_{-1}'\varepsilon)$.

We first discuss the limiting behavior of the different elements of the third expression of the above mentioned definition of $\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) \backsimeq G(\mathbf{\Pi}_{\mathrm{B}},\ \hat{\Omega}) - G(\mathbf{\Pi}_{\mathrm{A}},\ \hat{\Omega})$. The limiting behavior of F is such that

$$\frac{1}{T^2}F \Rightarrow$$
$$\left((I_N\otimes I_k)\otimes\begin{pmatrix}\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0\dots 0 & 0\\ 0 & \ddots & 0\\ 0 & 0\dots 0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\end{pmatrix}\right)$$
$$\left((I_N\otimes I_k)\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)$$
$$\left((I_N\otimes I_k)\otimes\begin{pmatrix}\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0\dots 0 & 0\\ 0 & \ddots & 0\\ 0 & 0\dots 0 & \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\end{pmatrix}'\right)$$
$$\begin{pmatrix}(e_1\otimes I_k)\otimes(e_1\otimes I_k) & \cdots & (e_N\otimes I_k)\otimes(e_N\otimes I_k)\end{pmatrix},$$

and given the limiting behavior of $(Y_{-1}'Y_{-1})^{-1}$ from Appendix B, $F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F$ converges as

$$\frac{1}{T^2}F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F\Rightarrow$$
$$\begin{pmatrix} (e_1\otimes I_k)'\otimes\left(e_1'\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\right) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}) \end{pmatrix}\left(\Omega^{-1}\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)$$
$$\begin{pmatrix} (e_1\otimes I_k)'\otimes\left(e_1'\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\right) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}) \end{pmatrix}=$$
$$\begin{pmatrix} I_k\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_k\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix}\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}$$
$$\left(\Omega^{-1}\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)$$
$$\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\begin{pmatrix} I_k\otimes\beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_k\otimes\beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix}'.$$

As a consequence, the inverse of this expression converges as

$$T^2\left(F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F\right)^{-1}\Rightarrow$$
$$\begin{pmatrix} I_k\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_k\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}$$
$$\left[\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}\left(\Omega^{-1}\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)\right.$$
$$\left.\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\right]^{-1}$$
$$\begin{pmatrix} I_k\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_k\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}'.$$

The limiting behavior of $F'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\mathrm{vec}(Y_{-1}'\varepsilon)$ is such that

$$
\begin{aligned}
&\frac{1}{T}F'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\mathrm{vec}(Y_{-1}'\varepsilon) \Rightarrow \\
&\qquad \begin{pmatrix} (e_1 \otimes I_k)' \otimes \left(e_1' \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1}\right) \\ \vdots \\ (e_N \otimes I_k)' \otimes \left(e_N' \otimes \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1}\right) \end{pmatrix} \mathrm{vec}\left(\Lambda^{\frac{1}{2}} \int B_{N(k-r)} dB_{Nk}' \Omega^{-\frac{1}{2}\prime}\right) = \\
&\begin{pmatrix} I_k \otimes \beta_{1\perp}(\alpha_{1\perp}'\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_k \otimes \beta_{N\perp}(\alpha_{N\perp}'\beta_{N\perp})^{-1} \end{pmatrix} \begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix} \\
&\qquad \mathrm{vec}\left(\Lambda^{\frac{1}{2}} \int B_{N(k-r)} dB_{Nk}' \Omega^{-\frac{1}{2}\prime}\right).
\end{aligned}
$$

Combining these results, we obtain that

$$
\begin{aligned}
&TH_\perp' \left(F'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})F\right)^{-1} F'(\hat{\Omega}^{-1} \otimes (Y_{-1}'Y_{-1})^{-1})\mathrm{vec}(Y_{-1}'\varepsilon) \Rightarrow \\
&\begin{pmatrix} \alpha_{1\perp}' \otimes \beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_{N\perp}' \otimes \beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix} \\
&\left[\begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix} \left(\Omega^{-1} \otimes \Lambda^{\frac{1}{2}} \left(\int B_{N(k-r)} B_{N(k-r)}'\right) \Lambda^{\frac{1}{2}\prime}\right) \right. \\
&\qquad \left. \begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix}' \right]^{-1} \\
&\qquad \begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix} \mathrm{vec}\left(\Lambda^{\frac{1}{2}} \int B_{N(k-r)} dB_{Nk}' \Omega^{-\frac{1}{2}\prime}\right).
\end{aligned}
$$

The matrix $H_\perp'\left(F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F\right)^{-1}H_\perp$ converges according to

$$
\begin{aligned}
&T^2 H_\perp'\left(F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})E\right)^{-1}H_\perp \Rightarrow \\
&\quad\begin{pmatrix} \alpha_{1\perp}'\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_{N\perp}'\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix} \\
&\quad\left[\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}\left(\Omega^{-1}\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)\right. \\
&\quad\left.\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\right]^{-1} \\
&\quad\begin{pmatrix} \alpha_{1\perp}'\otimes\beta_{1\perp}(\beta_{1\perp}'\beta_{1\perp})^{-1}\alpha_{1\perp}'\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_{N\perp}'\otimes\beta_{N\perp}(\beta_{N\perp}'\beta_{N\perp})^{-1}\alpha_{N\perp}'\beta_{N\perp} \end{pmatrix}',
\end{aligned}
$$

such that the limiting behavior of its inverse is characterized by

$$
\begin{aligned}
&\frac{1}{T^2}\left(H_\perp'\left(F'(\hat{\Omega}^{-1}\otimes(Y_{-1}'Y_{-1})^{-1})F\right)^{-1}H_\perp\right)^{-1} \Rightarrow \\
&\quad\begin{pmatrix} I_{k-r}\otimes\beta_{1\perp}\left(\alpha_{1\perp}'\beta_{1\perp}\right)^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_{k-r}\otimes\beta_{N\perp}\left(\alpha_{N\perp}'\beta_{N\perp}\right)^{-1} \end{pmatrix} \\
&\quad\left\{\begin{pmatrix} \alpha_{1\perp}'\otimes I_{k-r} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_{N\perp}'\otimes I_{k-r} \end{pmatrix}\right. \\
&\quad\left[\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}\left(\Omega^{-1}\otimes\Lambda^{\frac{1}{2}}\left(\int B_{N(k-r)}B_{N(k-r)}'\right)\Lambda^{\frac{1}{2}\prime}\right)\right. \\
&\quad\left.\left.\begin{pmatrix} (e_1\otimes I_k)'\otimes(e_1'\otimes I_{k-r}) \\ \vdots \\ (e_N\otimes I_k)'\otimes(e_N'\otimes I_{k-r}) \end{pmatrix}'\right]^{-1}\begin{pmatrix} \alpha_{1\perp}'\otimes I_{k-r} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_{N\perp}'\otimes I_{k-r} \end{pmatrix}'\right\}^{-1} \\
&\quad\begin{pmatrix} I_{k-r}\otimes\beta_{1\perp}\left(\alpha_{1\perp}'\beta_{1\perp}\right)^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_{k-r}\otimes\beta_{N\perp}\left(\alpha_{N\perp}'\beta_{N\perp}\right)^{-1} \end{pmatrix}'.
\end{aligned}
$$

Before we construct the limiting behavior of $\mathrm{LR}(\Pi_{\mathrm{B}}\,|\Pi_{\mathrm{A}}) \backsimeq G(\mathbf{\Pi}_{\mathrm{B}},\ \hat{\Omega}) - G(\mathbf{\Pi}_{\mathrm{A}},\ \hat{\Omega})$, we note that

$$\begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix} \left(\Omega^{-\frac{1}{2}} \otimes \Lambda^{\frac{1}{2}}\right) \left(\Omega^{-\frac{1}{2}} \otimes \Lambda^{\frac{1}{2}}\right)' \begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix}'$$
$$= \begin{pmatrix} \Sigma_1 \otimes \Theta_1 \\ \vdots \\ \Sigma_N \otimes \Theta_N \end{pmatrix} \begin{pmatrix} \Sigma_1 \otimes \Theta_1 \\ \vdots \\ \Sigma_N \otimes \Theta_N \end{pmatrix}'$$
$$= \begin{pmatrix} \left(\Omega^{-1}\right)_{11} \otimes \Lambda_{11} & \cdots & \left(\Omega^{-1}\right)_{1N} \otimes \Lambda_{1N} \\ \vdots & \ddots & \vdots \\ \left(\Omega^{-1}\right)_{N1} \otimes \Lambda_{N1} & \cdots & \left(\Omega^{-1}\right)_{NN} \otimes \Lambda_{NN} \end{pmatrix},$$

where $\Omega^{-\frac{1}{2}} = (\Sigma_1' \cdots \Sigma_N')'$ with Σ_i is $k \times Nk$ and $\Lambda^{\frac{1}{2}} = (\Theta_1' \cdots \Theta_N')'$ with Θ_i is $k \times Nk$, in all cases for $i = 1, \ldots, N$. We can define Ω^{-1} and Θ as:

$$\Omega^{-1} = \begin{pmatrix} \left(\Omega^{-1}\right)_{11} & \cdots & \left(\Omega^{-1}\right)_{1N} \\ \vdots & \ddots & \vdots \\ \left(\Omega^{-1}\right)_{N1} & \cdots & \left(\Omega^{-1}\right)_{NN} \end{pmatrix},\ \Lambda = \begin{pmatrix} \Lambda_{11} & \cdots & \Lambda_{1N} \\ \vdots & \ddots & \vdots \\ \Lambda_{N1} & \cdots & \Lambda_{NN} \end{pmatrix},$$

$\left(\Omega^{-1}\right)_{ij}$, Λ_{ij} is $k \times k$ and $i, j = 1, \ldots, N$. This covariance matrix can be specified, using a Choleski decomposition, as

$$\begin{pmatrix} \left(\Omega^{-1}\right)_{11} \otimes \Lambda_{11} & \cdots & \left(\Omega^{-1}\right)_{1N} \otimes \Lambda_{1N} \\ \vdots & \ddots & \vdots \\ \left(\Omega^{-1}\right)_{N1} \otimes \Lambda_{N1} & \cdots & \left(\Omega^{-1}\right)_{NN} \otimes \Lambda_{NN} \end{pmatrix} = AA',$$

where A is $Nk(k-r) \times Nk(k-r)$. It then results that

$$A^{-1} \begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix} \left(\Omega^{-\frac{1}{2}} \otimes \Lambda^{\frac{1}{2}}\right) \left(I_{Nk} \otimes \left(\int B_{N(k-r)} B_{N(k-r)}'\right)\right)$$
$$\times \left(\Omega^{-\frac{1}{2}} \otimes \Lambda^{\frac{1}{2}}\right)' \begin{pmatrix} (e_1 \otimes I_k)' \otimes (e_1' \otimes I_{k-r}) \\ \vdots \\ (e_N \otimes I_k)' \otimes (e_N' \otimes I_{k-r}) \end{pmatrix}' A^{-1\prime}$$
$$= A^{-1} \begin{pmatrix} \left(\Omega^{-1}\right)_{11} \otimes \Theta_1 \left(\int B_{N(k-r)} B_{N(k-r)}'\right) \Theta_1' & \cdots & \left(\Omega^{-1}\right)_{1N} \otimes \Theta_1 \left(\int B_{N(k-r)} B_{N(k-r)}'\right) \Theta_N' \\ \vdots & \ddots & \vdots \\ \left(\Omega^{-1}\right)_{N1} \otimes \Theta_N \left(\int B_{N(k-r)} B_{N(k-r)}'\right) \Theta_1' & \cdots & \left(\Omega^{-1}\right)_{NN} \otimes \Theta_N \left(\int B_{N(k-r)} B_{N(k-r)}'\right) \Theta_N' \end{pmatrix} A^{-1\prime}$$
$$= \begin{pmatrix} (I_k \otimes \left(\int B_{k-r,1} B_{k-r,1}'\right)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & (I_k \otimes \left(\int B_{k-r,N} B_{k-r,N}'\right)) \end{pmatrix},$$

where $B_{N(k-r)} = (B_{k-r,1}' \cdots B_{k-r,N}')'$ and $B_{k-r,i}$ is a $k-r$ dimensional Brownian motion with identity covariance matrix. Using the above, we can specify the limiting behavior of $\mathrm{LR}(\Pi_{\mathrm{B}}\,|\Pi_{\mathrm{A}}) \backsimeq G(\mathbf{\Pi}_{\mathrm{B}},\ \hat{\Omega}) -$

$G(\mathbf{\Pi}_{\mathrm{A}},\ \hat{\Omega})$ as

$$\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) \Rightarrow$$

$$\begin{pmatrix} \mathrm{vec}(\Theta_1 \int B_{N(k-r)} dB'_{Nk}\Sigma'_1) \\ \vdots \\ \mathrm{vec}(\Theta_N \int B_{N(k-r)} dB'_{Nk}\Sigma'_N) \end{pmatrix}' A^{-1\prime}$$

$$\left[\begin{pmatrix} (I_k \otimes \left(\int B_{k-r,1}B'_{k-r,1}\right)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & (I_k \otimes \left(\int B_{k-r,N}B'_{k-r,N}\right)) \end{pmatrix}\right]^{-1}$$

$$A^{-1}\begin{pmatrix} \alpha'_{1\perp}\otimes\beta_{1\perp}(\beta'_{1\perp}\beta_{1\perp})^{-1}\alpha'_{1\perp}\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha'_{N\perp}\otimes\beta_{N\perp}(\beta'_{N\perp}\beta_{N\perp})^{-1}\alpha'_{N\perp}\beta_{N\perp} \end{pmatrix}'$$

$$\begin{pmatrix} I_{k-r}\otimes\beta_{1\perp}(\alpha'_{1\perp}\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_{k-r}\otimes\beta_{N\perp}(\alpha'_{N\perp}\beta_{N\perp})^{-1} \end{pmatrix}$$

$$\left\{\begin{pmatrix} \alpha'_{1\perp}\otimes I_{k-r} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha'_{N\perp}\otimes I_{k-r} \end{pmatrix} A^{-1\prime}\right.$$

$$\left[\begin{pmatrix} (I_k \otimes \left(\int B_{k-r,1}B'_{k-r,1}\right)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & (I_k \otimes \left(\int B_{k-r,N}B'_{k-r,N}\right)) \end{pmatrix}\right]^{-1}$$

$$\left. A^{-1}\begin{pmatrix} \alpha'_{1\perp}\otimes I_{k-r} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha'_{N\perp}\otimes I_{k-r} \end{pmatrix}'\right\}^{-1}$$

$$\begin{pmatrix} I_{k-r}\otimes\beta_{1\perp}(\alpha'_{1\perp}\beta_{1\perp})^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_{k-r}\otimes\beta_{N\perp}(\alpha'_{N\perp}\beta_{N\perp})^{-1} \end{pmatrix}'$$

$$\begin{pmatrix} \alpha'_{1\perp}\otimes\beta_{1\perp}(\beta'_{1\perp}\beta_{1\perp})^{-1}\alpha'_{1\perp}\beta_{1\perp} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha'_{N\perp}\otimes\beta_{N\perp}(\beta'_{N\perp}\beta_{N\perp})^{-1}\alpha'_{N\perp}\beta_{N\perp} \end{pmatrix} A^{-1\prime}$$

$$\left[\begin{pmatrix} (I_k \otimes \left(\int B_{k-r,1}B'_{k-r,1}\right)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & (I_k \otimes \left(\int B_{k-r,N}B'_{k-r,N}\right)) \end{pmatrix}\right]^{-1} A^{-1}\begin{pmatrix} \mathrm{vec}(\Theta_1 \int B_{N(k-r)} dB'_{Nk}\Sigma'_1) \\ \vdots \\ \mathrm{vec}(\Theta_N \int B_{N(k-r)} dB'_{Nk}\Sigma'_N) \end{pmatrix}.$$

We note that

$$A^{-1}\begin{pmatrix} \mathrm{vec}(\Theta_1 \int B_{N(k-r)} dB'_{Nk}\Sigma'_1) \\ \vdots \\ \mathrm{vec}(\Theta_N \int B_{N(k-r)} dB'_{Nk}\Sigma'_N) \end{pmatrix} = \begin{pmatrix} \mathrm{vec}(\int B_{k-r,1} dB'_{k,1}) \\ \vdots \\ \mathrm{vec}(\int B_{k-r,N} dB'_{k,N}) \end{pmatrix},$$

where $B_{Nk} = (B'_{k,1} \cdots B'_{k,N})'$, $B_{k,i}$ is a k dimensional Brownian motion with identity covariance matrix for individual i. Using the above, we can simplify the limiting behavior of $\mathrm{LR}(\Pi_{\mathrm{B}}\,|\Pi_{\mathrm{A}}) \backsimeq G(\mathbf{\Pi}_{\mathrm{B}},\ \hat{\Omega}) - G(\mathbf{\Pi}_{\mathrm{A}},\ \hat{\Omega})$ as

$$\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) \Rightarrow$$

$$\begin{pmatrix} \mathrm{vec}(\left(\int B_{k-r,1}B'_{k-r,1}\right)^{-1} \int B_{k-r,1} dB'_{k,1}) \\ \vdots \\ \mathrm{vec}(\left(\int B_{k-r,N}B'_{k-r,N}\right)^{-1} \int B_{k-r,1} dB'_{k,N}) \end{pmatrix}' P$$

$$\left\{ P' \left[\begin{pmatrix} (I_k \otimes \left(\int B_{k-r,1}B'_{k-r,1}\right)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & (I_k \otimes \left(\int B_{k-r,N}B'_{k-r,N}\right)) \end{pmatrix} \right]^{-1} P \right\}^{-1}$$

$$P' \begin{pmatrix} \mathrm{vec}(\left(\int B_{k-r,1}B'_{k-r,1}\right)^{-1} \int B_{k-r,1} dB'_{k,1}) \\ \vdots \\ \mathrm{vec}(\left(\int B_{k-r,N}B'_{k-r,N}\right)^{-1} \int B_{k-r,N} dB'_{k,N}) \end{pmatrix},$$

where $P = A^{-1}\begin{pmatrix} \alpha_{1\perp}\otimes I_{k-r} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_{N\perp}\otimes I_{k-r} \end{pmatrix}$ and P is $Nk(k-r)\times N(k-r)^2$. We can now further simplify the above expression as P is a scaling parameter that essentially only affects the sizes of the Brownian motion increments $dB_{k,i}$ and the I_k matrices. Using this, the limiting behavior of $\mathrm{LR}(\Pi_{\mathrm{B}}\,|\Pi_{\mathrm{A}})$ can be further simplified to

$$\mathrm{LR}(\Pi_{\mathrm{B}}|\Pi_{\mathrm{A}}) \Rightarrow$$

$$\begin{pmatrix} \mathrm{vec}(\int B_{k-r,1} dB'_{k-r,1}) \\ \vdots \\ \mathrm{vec}(\int B_{k-r,N} dB'_{k-r,N}) \end{pmatrix}' \begin{pmatrix} \left(\int B_{k-r,1}B'_{k-r,1}\right)^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \left(\int B_{k-r,N}B'_{k-r,N}\right)^{-1} \end{pmatrix}$$

$$\begin{pmatrix} \mathrm{vec}(\int B_{k-r,1} dB'_{k-r,1}) \\ \vdots \\ \mathrm{vec}(\int B_{k-r,N} dB'_{k-r,N}) \end{pmatrix} =$$

$$\sum_{i=1}^{N} tr\left(\left(\int B_{k-r,i} dB'_{k-r,i}\right)' \left(\int B_{k-r,i}B'_{k-r,i}\right)^{-1} \left(\int B_{k-r,i} dB'_{k-r,i}\right)\right),$$

which is the sum of the limiting behavior of N Johansen trace statistics.

References

Billingsley, P., 1986, *Probability and Measure*, Wiley, New York.

Engle, R.F. and C.W.J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation and Testing, *Econometrica* **55**, 251–276.

Groen, J.J.J., 1999, The Monetary Exchange Rate Model as a Long-Run Phenomenon. Revision conditionally accepted for *Journal of International Economics*.

Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* **50**, 1029–1054.

Johansen, S., 1991, Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica* **59**, 1551–1580.

Johansen, S., 1996, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, 2nd edition, Oxford University Press, Oxford.

Kao, C., 1999, Spurious Regression and Residuals-Based Tests for Cointegration in Panel Data, *Journal of Econometrics* **90**, 1–44.

Kleibergen, F., 1999, Reduced Rank Regression Using Generalized Method of Moments Estimators, *in* L. Matyas (editor), *Generalised Method of Moments Estimation*, Cambridge University Press, Cambridge.

Kleibergen, F., J.-P. Urbain and H.K. Van Dijk, 1999, Oil Price Shocks and Long Run Price and Import Demand Behavior, *Annals of the Institute of Mathematical Statistics* , forthcoming.

Larsson, R., J. Lyhagen and M. Löthgren, 1998, Likelihood-Based Cointegration Tests in Heterogeneous Panels, *Working Paper in Economics and Finance 250*, Stockholm School of Economics.

Levin, A. and Chien-Fu Lin, 1992, Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, *Discussion Paper 92-23*, Department of Economics, University of California at San Diego.

Lütkepohl, H., 1993, *Introduction to Multiple Time Series Analysis*, 2nd edition, Springer-Verlag, Berlin.

Mussa, M., 1976, The Exchange Rate, the Balance of Payments and Monetary and Fiscal Policy under a Regime of Controlled Floating, *Scandinavian Journal of Economics* **78**, 229–248.

O'Connell, P., 1998, The Overvaluation of Purchasing Power Parity, *Journal of International Economics* **44**, 1–19.

Osterwald-Lenum, M., 1992, A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics, *Oxford Bulletin of Economics and Statistics* **54**, 461–471.

Papell, D.H., 1997, Searching for Stationarity: Purchasing Power Parity Under the Current Float, *Journal of International Economics* **43**, 313–332.

Park, J.Y., 1992, Canonical Cointegrating Regressions, *Econometrica* **60**, 119–143.

Park, J.Y. and M. Ogaki, 1991, Seemingly Unrelated Canonical Cointegrating Regressions, *Working Paper 280*, Rochester Center for Economic Research, University of Rochester.

Pedroni, P., 1995, Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis, *Working Paper 95-013*, Department of Economics, Indiana University.

Pedroni, P., 1996, Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity, *Working Paper 96-020*, Department of Economics, Indiana University.

Pesaran, M.H., Y. Sin and R.P. Smith, 1998, Pooled Mean Group Estimation of Dynamic Heterogeneous Panels, *Journal of the American Statistical Association*, forthcoming.

Phillips, P.C.B., 1991, Optimal Inference in Cointegrated Systems, *Econometrica* **59**, 283–306.

Phillips, P.C.B., 1995, Fully Modified Least Squares and Vector Autoregression, *Econometrica* **63**, 1023–1078.

Sarantis, N., 1994, The Monetary Exchange Rate Model in the Long Run: An Empirical Investigation, *Weltwirtschaftliches Archiv* **130**, 698–711.

Shiller, R. and P. Perron, 1985, Testing the Random Walk Hypothesis: Power versus Frequency of Observation, *Economics Letters* **18**, 381–386.

Toda, H.Y. and P.C.B. Phillips, 1993, Vector Autoregressions and Causality, *Econometrica* **61**, 1367–1393.