EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 3, No. 3, 2010, 541-571
ISSN 1307-5543 – www.ejpam.com

SPECIAL ISSUE ON
GRANGER ECONOMETRICS AND STATISTICAL MODELING
DEDICATED TO THE MEMORY OF PROF. SIR CLIVE W.J. GRANGER

Likelihood Ratio Tests on Cointegrating Vectors, Disequilibrium Adjustment Vectors, and Their Orthogonal Complements

Norman Morin*

Division of Research and Statistics, Federal Reserve Board, Washington, DC, USA

Abstract. Cointegration theory provides a flexible class of statistical models that combine long-run (cointegrating) relationships and short-run dynamics. This paper presents three likelihood ratio (LR) tests for simultaneously testing restrictions on cointegrating relationships and on how quickly each variable in the system reacts to the deviation from equilibrium implied by the cointegrating relationships. Both the orthogonal complements of the cointegrating vectors and of the vectors of adjustment speeds have been used to define the common stochastic trends of a nonstationary system. The restrictions implicitly placed on the orthogonal complements of the cointegrating vectors and of the adjustment speeds are identified for a class of LR tests, including those developed in this paper. It is shown how these tests can be interpreted as tests for restrictions on the orthogonal complements of the cointegrating relationships and of their adjustment vectors, which allow one to combine and test for economically meaningful restrictions on cointegrating relationships and on common stochastic trends.

2000 Mathematics Subject Classifications: 62H12, 62H15

Key Words and Phrases: Cointegration, common stochastic trend, likelihood ratio tests

1. Introduction

Since its introduction by Granger [14,15] cointegration has become a widely investigated and extensively used tool in multivariate time series analysis. Cointegrated

Email address: norman.j.morin@frb.gov

*The opinions expressed are those of the author and not necessarily those of the Board of Governors of the Federal Reserve System or its staff.

models combine short-run dynamics and long-run relationships in a framework that lends itself to investigating these features in economic data. The relationship between cointegrated systems, their vector autoregressive (VAR) and vector moving-average representations, and vector error-correction models (VECM) were developed by Granger in [14,15] and by Engle and Granger in [7].

In a cointegrated system of time series, the cointegrating vectors can be interpreted as the long-run equilibrium relationships among the variables towards which the system will tend to be drawn. Economic theories and economic models may imply long-run relationships among variables. Certain ratios or spreads between nonstationary variables are expected to be stationary, that is, these variables are cointegrated with given cointegrating vectors. For example, neoclassical growth models imply "balanced growth" among income, consumption, and investment (for example [29, 41]), implying that their ratios are mean-reverting. Other theories, rather than implying given ratios or spreads are cointegrated, may imply that *some* linear combinations of the variables are stationary, that is, the variables are cointegrated without specifying the cointegrating relationships (for example [25]).

Johansen's maximum likelihood approach to cointegrated models [19] provides an efficient procedure for the estimation of cointegrated systems and provides a useful framework in which to test restrictions of the sorts mentioned above. For example, Johansen [19, 21] and Johansen and Juselius [25, 26] derive likelihood ratio tests for various structural hypotheses concerning the cointegrating relationships and the speed of adjustment to the disequilibrium implied by the cointegrating relationships (or weights); Konishi and Granger [30] use this approach to derive and test for separation cointegration, and Gonzalo and Granger [12] use this framework for estimation of and testing for their multivariate version of Quah's [37] permanent and transitory (P-T) decomposition.

Further, building on the univariate work of Beveridge and Nelson [1] and the multivariate generalization by Stock and Watson [42], cointegration analysis may be used to decompose a system of variables into permanent components (based on the variables' common stochastic trends) and temporary (or cyclical) components. Several methods have been proposed to separate cointegrated systems into their permanent and temporary components (for example, [12, 21, and 27]). In each case, the permanent component is based either on the orthogonal complements of the cointegrating relationships or on the orthogonal complements of the disequilibrium adjustments to the cointegrating relationships.

In this paper, new hypothesis tests are presented in Johansen's maximum likelihood framework that allow one to combine restrictions on the cointegrating relationships and on their disequilibrium adjustments. These tests possess closed-form solutions and do not require iterative methods to estimate the restricted parameters under the null hypothesis. Secondly, both for Johansen's likelihood ratio tests for coefficient restrictions and for the new tests presented below, the restrictions implicitly placed on the orthogonal complements of the cointegrating relationships and on the orthogonal complements of the adjustment speeds are presented. Johansen's tests and the tests developed in this paper can be interpreted as tests of restrictions on the various definitions of common stochastic trends, since these definitions depend on the orthogonal complements either of the cointegrating relationships or of the disequilibrium adjustments. Thus, one has great flexibility in formulating and testing hypotheses of economic interest simultaneously on the cointegrating relationships and on the common stochastic trends—the long-run relationships among the variables in the system and the variables driving the trending behavior the system, respectively.

The organization of this paper is as follows: In section 2, the basic model and notation are introduced, and maximum likelihood estimation of the unrestricted model is briefly described. In section 3, likelihood ratio tests for restrictions on cointegrating relationships and on their weights are briefly described, and three new tests in this framework are presented. In section 4, the implications for the orthogonal complements of the cointegrating vectors and of the adjustment vectors are developed for the tests described in section 3. It is shown how these tests can be used for testing restrictions on the orthogonal complements of cointegrating vectors and on the orthogonal complements of the disequilibrium adjustment vectors—thus allowing for combinations of tests on cointegrating relationships and on the different definitions of common stochastic trends. Section 5 concludes, and the appendix contains the mathematical proofs.

2. The Unrestricted Cointegrated Model

Let $I(d)$ denote a time series that is integrated of order *d*, that is, *d* applications of the differencing filter, $\Delta = 1 - L$, yield a stationary process. Let X_t be a *p×1* vector of possibly *I*(1) time series defined by the k^{th}-order vector autoregression (VAR),

$$X_t = \sum_{i=1}^{k} \Pi_i X_{t-i} + \Phi D_t + \varepsilon_t \ , \ t = 1, \ldots, T \ , \tag{1}$$

and generated by initial values $X_{-k}, \ldots, X_0$, by *p*-dimensional normally-distributed zero-mean random variables $\{\varepsilon_t\}_{t=0}^{T}$ with variance matrix Ω and by a vector of deterministic components D_t (possibly constants, linear trends, and seasonal and other dummy variables). Using the lag polynomial expression for (1),

$$\Pi(L) X_t = \Phi D_t + \varepsilon_t \ , \tag{2}$$

where $\Pi(L) = I - \sum_{i=1}^{k} \Pi_i L^i$, the VAR in the levels in (1) can be rewritten in first differences as

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t \ , \tag{3}$$

where $\Pi = -\Pi(1) = -\left(I - \sum_{i=1}^{k} \Pi_i \right)$ and $\Gamma_i = -\sum_{j=i+1}^{k} \Pi_j$, $i = 1, \ldots, k-1$.

The long-run behavior of the system depends on the rank of the *p×p* matrix Π. If the matrix has rank 0 (that is, $\Pi = 0$) then there are *p* unit roots in the system, and (3) is simply a traditional VAR in differences. If Π has full rank *p*, then X_t is an I(0) process, that is, X_t is stationary in its levels. If the rank of Π is *r* with $0 < r < p$, then X_t is said to be cointegrated of order *r*. This implies that there are $r < p$ linear combinations of X_t that are stationary. Granger's Representation Theorem [7] shows that if X_t is cointegrated of order *r* (the *p×p* matrix Π has rank *r*), one can write $\Pi = \alpha\beta'$, where both α and β are *p×r* matrices of full column rank. This and some fairly general assumptions about initial distributions allow one to write (1) as the vector error-correction model (VECM):

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t \,. \tag{4}$$

The matrix β contains the r cointegrating vectors, and $\beta' X_t$ are the r stationary linear combinations of X_t. The matrix β can be interpreted as r equilibrium relationships among the variables, and the difference between the current value of the r cointegrating relationships, $\beta' X_t$, and their expected values can be interpreted as measures of disequilibrium from the r different long-run relationships. The matrix α in (4) measures how quickly ΔX_t reacts to the deviation from equilibrium implied by $\beta' X_t$.

Given a $p \times r$ matrix of full column rank, A, an orthogonal complement of A, denoted $A_\perp$, is a $p \times (p\text{-}r)$ matrix of full column rank such that $A_\perp' A = 0$. It is often necessary to calculate the orthogonal complements of β and α in order to form the p-r common I(1) stochastic trends of a cointegrated system; for example, Gonzalo and Granger [12] propose $\alpha_\perp' X_t$ as the common stochastic trends and $\beta_\perp \left(\alpha_\perp' \beta_\perp\right)^{-1} \alpha_\perp' X_t$ as the permanent components for a cointegrated system; Johansen [21] proposes the random walks $\alpha_\perp' \Gamma(L) X_t$ as a cointegrated system's common stochastic trends and $\beta_\perp \left(\alpha_\perp' \Gamma(1) \beta_\perp\right)^{-1} \alpha_\perp' \Gamma(L) X_t$ as its permanent components.

Several methods have been proposed for identifying, estimating, and conducting inference in a cointegrated system (see [11] and [45] for explanations of several methods and evaluations of their properties). This paper uses the efficient maximum likelihood framework of Johansen [19]. The log-likelihood function for the parameters in (4) is

$$\begin{aligned} \log L\left(\alpha, \beta, \Omega, \Gamma_1, \ldots, \Gamma_{k-1}, \Phi\right) = & -\frac{Tp}{2} \log 2\pi - \frac{T}{2} \log \left|\Omega\right| \ldots \\ & -\frac{1}{2} \sum_{t=1}^{T} \left[\left(\Delta X_t - \alpha\beta' X_{t-1} - \textstyle\sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} - \Phi D_t \right)' \ldots \right. \\ & \left. \times \Omega^{-1} \left(\Delta X_t - \alpha\beta' X_{t-1} - \textstyle\sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} - \Phi D_t \right) \right] \end{aligned} \tag{5}$$

.

Maximum likelihood estimation of the parameters in (5) involves successively concentrating the likelihood function until it is a function solely of β. To do this one forms two sets of $p \times 1$ residual vectors, R_{0t} and R_{1t}, by regressing, in turn, ΔX_t and X_{t-1} on k-1 lags of ΔX_t and the deterministic components.

The VECM in (4) can then be written as

$$R_{0t} = \alpha\beta' R_{1t} + \varepsilon_t, \; t = 1, \ldots, T \,. \tag{6}$$

This equation is the basis from which one derives the hypothesis tests on the cointegrating vectors β, on the disequilibrium adjustment parameters α, and on their orthogonal complements, $\beta_\perp$ and $\alpha_\perp$. The equation (6) has two unknown parameter matrices, α and β. Maximizing the likelihood function is equivalent to estimating the parameters in (6) via reduced rank regression methods. Since this involves the product of two unknown full-column rank matrices in (6), estimating these parameters requires solving an eigenvalue problem.

Defining the moment matrices for the residual series,

$$S_{ij} = \frac{1}{T}\sum_{t=1}^{T} R_{it}R_{ij}', \; i, j = 0,1, \tag{7}$$

for a given set of cointegrating vectors, β, one estimates the adjustment parameters, α,□ by regressing R_{0t} on $\beta' R_{1t}$ to get

$$\hat{\alpha}(\beta) = S_{01}\beta\left(\beta' S_{11}\beta\right)^{-1}. \tag{8}$$

The maximum likelihood estimator for the residual variance-covariance matrix is

$$\hat{\Omega}(\beta) = S_{00} - S_{01}\beta\left(\beta' S_{11}\beta\right)^{-1}\beta' S_{10}. \tag{9}$$

As shown in [19], one may write the likelihood function, apart from a constant, as

$$L(\beta)_{\max}^{-2/T} = \left|\hat{\Omega}(\beta)\right|, \tag{10}$$

which can be expressed as a function of $\hat{\beta}$,

$$L\left(\hat{\beta}\right)_{\max}^{-2/T} = \frac{\left|S_{00}\right|\left|\hat{\beta}' S_{11}\hat{\beta} - \hat{\beta}' S_{10}S_{00}^{-1}S_{01}\hat{\beta}\right|}{\left|\hat{\beta}' S_{11}\hat{\beta}\right|}. \tag{11}$$

As shown in [19], maximizing the likelihood function with respect to β is equivalent to minimizing (11), which is accomplished by solving the eigenvalue problem

$$\left|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right| = 0 \tag{12}$$

for eigenvalues $1 > \hat{\lambda}_1 > \ldots > \hat{\lambda}_p$ and corresponding eigenvectors $\hat{V} = \left(\hat{v}_1, \ldots, \hat{v}_p\right)$ normalized by $\hat{V}' S_{11}\hat{V} = I_p$. Thus the maximum likelihood estimate for the cointegrating vectors β is

$$\hat{\beta} = \left(\hat{v}_1, \ldots, \hat{v}_r\right), \tag{13}$$

and the normalization implies that the estimate of the weights in (8) is

$$\hat{\alpha} = S_{01}\hat{\beta}. \tag{14}$$

Then, apart from a constant, the maximized likelihood can be written as

$$L_{\max}^{-2/T} = \left|S_{00}\right|\prod_{i=1}^{r}\left(1 - \hat{\lambda}_i\right). \tag{15}$$

Likelihood ratio tests of the hypothesis of *r* unrestricted cointegrating relationships in the unrestricted VAR model and for *r* unrestricted cointegrating relationships against the alternative of *r+1* unrestricted cointegrating relationships—the trace and maximum eigenvalue tests—are derived in [19]. The asymptotic distribution of the trace and maximum eigenvalue tests for different deterministic components may be found in [19] and [25], and the tabulated critical values for various values of *r* and for different deterministic components may be found in [20, 23, 34]; small-sample adjustments to the critical values that are based on response surface regressions may be found in [2] and [31].

The unrestricted orthogonal complements of β and α, $\beta_\perp$ and $\alpha_\perp$, can be estimated three ways: One may use the eigenvectors associated with the zero eigenvalues of $\beta\beta'$ and $\alpha\alpha'$ [12] (given a $p \times r$ matrix of full column rank *A*, one can quickly construct $A_\perp$ as the ordered eigenvectors corresponding to the *p-r* zero-eigenvalues of AA'), and one may estimate $\alpha_\perp$ as the eigenvectors corresponding to the *p-r* smallest eigenvalues that solve the dual of the eigenvalue problem in (12), $\left|\lambda S_{00} - S_{01}S_{11}^{-1}S_{10}\right| = 0$, normalized such that

$\hat{\alpha}_{\perp}' S_{00} \hat{\alpha}_{\perp} = I_{p-r}$, and by setting $\hat{\beta}_{\perp} = S_{10} \hat{\alpha}_{\perp}$. Johansen [23] shows one may estimate them from (12) by $S_{11}\left(v_{r+1}, \ldots, v_p\right)$ and $S_{00}^{-1} S_{01}\left(v_{r+1}, \ldots, v_p\right)$, respectively.

3. Testing Restrictions on β and α

Economic theory may suggest that certain ratios or spreads between variables will be cointegrating relationships. For example, some neoclassical growth models with a stochastic productivity shock imply "balanced growth" among income, consumption, and investment (that is, the ratios are cointegrated), and certain one-factor models of the term structure of the interest rates imply that the spreads between the different interest rate maturities will be cointegrated. One might also be interested in testing for the absence of certain variables in the system from any of the cointegrating relationships. Complicated restrictions on β or α may be formulated, for example, neutrality hypotheses in Mosconi and Giannini [32] and separation cointegration in Konishi and Granger [30]. Based on their maximum likelihood framework, Johansen [19, 21] and Johansen and Juselius [25, 26] formulate a series of likelihood ratio tests for linear restrictions on β or α and tests for a subset of known vectors in β or α. After briefly summarizing this set of five tests, three new tests for combining linear restrictions and known vectors will be derived.

The tests for restrictions on the cointegrating relationships and disequilibrium adjustment vectors described below are asymptotically chi-squared distributed. The finite sample properties of some of the tests have been studied (see, for example [18]) and are shown to have significant size distortions in small samples, though they generally perform well with larger samples. Johansen [24] introduces a Bartlett-type correction for tests (1) and (2) below that depend on the size of the system, the number of cointegrating vectors, the lag length in the VECM, the number of deterministic terms (restricted versus unrestricted), the parameter values, and the sample size under the null hypothesis. Haug [18] demonstrates that the Bartlett correction is successful in moving the empirical size of the test close to the nominal size of the test and also demonstrates that the power of the tests for restrictions on β depend on the speed of adjustment to the long-run equilibrium relationships in the system, with slower adjustment speeds leading to tests with lower power.

The tests below are all based on the reduced rank regression representation of the VECM in (4),

$$R_{0t} = \alpha \beta' R_{1t} + \varepsilon_t, \; t = 1, \ldots, T, \tag{16}$$

the same equation that is the starting point for the maximum likelihood estimates of the parameters of the VECM. The estimators and test statistics are all calculated in terms of the residual product moment matrices S_{ij}, $i, j = 0,1$ and by their eigenvalues. The parameter estimates under the restrictions and the maximized likelihood functions can be explicitly calculated; other tests not discussed here may be solved using iterative methods (see [6] and [22]). Denote the unrestricted model of at most *r* cointegrating relationships in the VECM (4) as $H(r)$. For any rectangular matrix with full column rank, *A*, define the notation $\overline{A} \equiv A\left(A'A\right)^{-1}$, which implies $A'\overline{A} = \overline{A}'A = I_r$. Five tests for restrictions on β and α from Johansen [19, 20] and Johansen and Juselius [25] are briefly described before turning to three new tests for restrictions on β and α.

(1) $H_0 : \beta = H\phi$ (Johansen [19]), (17)

where H $p \times r$ is known and ϕ $s \times r$ is unknown, $r \leq s < p$.

This test places the same p-s linear restrictions on all the vectors in β. The likelihood ratio test of H_0 in $H(r)$ is asymptotically distributed as χ^2 with $r(p\text{-}s)$ degrees of freedom. One can also use this test also to determine if a subset of the p variables do not enter the cointegrating relationships.

(2) $H_0 : \beta = [H, \theta]$ (Johansen and Juselius [25]), (18)

where H $p \times s$ is known, and θ $p \times (r\text{-}s)$ is unknown where $\theta = \bar{H}_\perp \phi$ with $\bar{H}_\perp$ $p \times (p\text{-}s)$ known and ϕ $(p\text{-}s) \times (r\text{-}s)$ unknown.

This test assumes s known cointegrating vectors and restricts the remaining r-s unknown cointegrating vectors to be orthogonal to them. The likelihood ratio test of H_0 in $H(r)$ is asymptotically distributed as χ^2 with $s(p\text{-}r)$ degrees of freedom.

(3) $H_0 : \alpha = A\psi$ (Johansen and Juselius [25]), (19)

where A $p \times m$ is known and ψ $m \times r$ is unknown, $m \leq r < p$.

This test places the same p-m linear restrictions on all disequilibrium adjustment vectors in α. This can be interpreted as a test of $B'\alpha = 0$ for $B = A_\perp$. The likelihood ratio test of H_0 in $H(r)$ is asymptotically distributed as χ^2 with $r(p\text{-}m)$ degrees of freedom. One may use (3) to test that some or all of the cointegrating relationships do not appear in the short run equation for a subset of the variables in the system, that is, that a subset of the variables do not error correct to some or all of the stochastic trends in the system.

(4) $H_0 : \alpha = [A, \tau]$ (Johansen [20]), (20)

where A $p \times m$ is known, and τ $p \times (r\text{-}m)$ is unknown where $\tau = \bar{A}_\perp \psi$ with $\bar{A}_\perp$ $p \times (p\text{-}m)$ known and ψ $(p\text{-}m) \times (r\text{-}m)$ unknown.

This test allows for m known adjustment vectors and restricts the remaining r-m adjustment vectors to be orthogonal to them. The likelihood ratio test of H_0 in $H(r)$ is asymptotically distributed as χ^2 with $m(p\text{-}r)$ degrees of freedom.

(5) $H_0 : \beta = H\phi,\ \alpha = A\psi$ (Johansen and Juselius [25]), (21)

where H $p \times s$, A $p \times m$ are known and ϕ $s \times r$, ψ $m \times r$ are unknown, $r \leq s < p$ and $r \leq m < p$.

This test combines tests (1) and (3), testing for cointegrating vectors with p-s common linear restrictions and adjustment vectors with p-m common linear restrictions. The likelihood ratio test of H_0 in $H(r)$ is asymptotically distributed as χ^2 with $r(p\text{-}s)+r(p\text{-}m)$ degrees of freedom.

In the same framework as the tests above, three new tests for simultaneous restrictions on β and α are presented.

(6) $H_0 : \beta = [H, \theta],\ \alpha = A\psi$ (22)

where H $p\times s$, A $p\times m$ are known; θ $p\times(r\text{-}s)$ is unknown where $\theta = \bar{H}_\perp \phi$ with $\bar{H}_\perp$ $p\times(p\text{-}s)$ known and ϕ $(p\text{-}s)\times(r\text{-}s)$ unknown; and ψ $m\times r$ is unknown, $s\le r\le m<p$.

This test combines tests (2) and (3), that is, it tests the restriction that s of the cointegrating vectors are known—restricting the remaining r-s cointegrating vectors to be orthogonal to them—and that the adjustment vectors share p-m linear restrictions. For example, if a system of variables includes a short-term and a long-term interest rate, (6) could be used to test whether the spread between the long-term and short-term interest rates was a cointegrating relationship and to test simultaneously whether the short-term interest rate failed to react to any of the cointegrating relationships in the system.

To calculate the test statistic and the estimated cointegrating relationships and adjustment vectors, the reduced rank regression (16) first is split into

$$\begin{aligned} \bar{A}' R_{0t} &= \psi_1 H' R_{1t} + \psi_2 \phi' \bar{H}_\perp' R_{1t} + \bar{A}' \varepsilon_t \\ A_\perp' R_{0t} &= A_\perp' \varepsilon_t \end{aligned}, \tag{23}$$

where ψ is partitioned conformably with β as $[\psi_1, \psi_2]$. In order to derive the test statistic and to estimate the restricted parameters under this hypothesis it is necessary to transform the product moment matrices, S_{ij}. Define two set of moment matrices:

$$S_{ij.A_\perp} = S_{ij} - S_{i0} A_\perp \left(A_\perp' S_{00} A_\perp \right)^{-1} A_\perp' S_{0j},\ i, j = 0,1 \tag{24}$$

and

$$S_{ij.A_\perp.H} = S_{ij.A_\perp} - S_{i1.A_\perp} H \left(H' S_{11.A_\perp} H \right)^{-1} H' S_{1j.A_\perp},\ i, j = 0,1. \tag{25}$$

The restricted estimators and the likelihood ratio test statistic and its asymptotic distribution are summarized in the following theorems.

Theorem 1. *Under the hypothesis* $H_0 : \beta = [H, \theta],\ \alpha = A\psi$ *where* H $p\times s$, A $p\times m$ *are known;* θ $p\times(r\text{-}s)$ *is unknown where* $\theta = \bar{H}_\perp \phi$ *with* $\bar{H}_\perp$ $p\times(p\text{-}s)$ *known and* ϕ $(p\text{-}s)\times(r\text{-}s)$ *unknown; and* ψ $m\times r$ *is unknown,* $s\le r\le m<p$*; the maximum likelihood estimators are found by the following steps:*

Solve the eigenvalue problem

$$\left| \lambda \bar{H}_\perp' S_{11.A_\perp.H} \bar{H}_\perp - \bar{H}_\perp' S_{10.A_\perp.H} A \left(A' S_{00.A_\perp.H} A \right)^{-1} A' S_{01.A_\perp.H} \bar{H}_\perp \right| = 0 \tag{26}$$

for eigenvalues $1 \ge \tilde{\lambda}_1 \ge \ldots \ge \tilde{\lambda}_{p-s} \ge 0$ *and corresponding eigenvectors* $\tilde{V} = \left(\tilde{v}_1, \ldots, \tilde{v}_{p-s} \right)$*, normalized so that* $\tilde{V}' \bar{H}_\perp' S_{11.A_\perp.H} \bar{H}_\perp \tilde{V} = I_{p-s}$*; and solve the eigenvalue problem*

$$\left| \rho H' S_{11.A_\perp} H - H' S_{10.A_\perp} A \left(A' S_{00.A_\perp} A \right)^{-1} A' S_{01.A_\perp} H \right| = 0 \tag{27}$$

for eigenvalues $1 \ge \tilde{\rho}_1 \ge \ldots \ge \tilde{\rho}_s \ge 0$.

Then the restricted estimators are

$$\hat{\phi} = \left(\tilde{v}_1, \ldots, \tilde{v}_{r-s} \right) \tag{28}$$

$$\hat{\beta} = \left[\hat{\beta}_1, \hat{\beta}_2 \right] = \left[H, \hat{\theta} \right] = \left[H, \bar{H}_\perp \hat{\phi} \right] \tag{29}$$

$$\hat{\psi}_2 = \bar{A}' S_{01.A_\perp.H} \bar{H}_\perp \hat{\phi} \tag{30}$$

$$\hat{\psi}_1 = \left(\bar{A}' S_{01.A_\perp} H - \hat{\psi}_2 \hat{\phi}' \bar{H}'_\perp S_{11.A_\perp} H\right)\left(H' S_{11.A_\perp} H\right)^{-1} \tag{31}$$

$$\begin{aligned}\hat{\alpha} = \left[A\hat{\psi}_1, A\hat{\psi}_2\right] = \Big[& A\left(A'A\right)^{-1} A'\left(S_{01.A_\perp}\hat{\beta}_1 - S_{01.A_\perp.H}\hat{\beta}_2\hat{\beta}_2' S_{11.A_\perp}\hat{\beta}_1\right)\left(\hat{\beta}_1' S_{11.A_\perp}\hat{\beta}_1\right)^{-1}, \ldots \\ & A\left(A'A\right)^{-1} A' S_{01.A_\perp.H}\hat{\beta}_2 \Big]\end{aligned} \tag{32}$$

and the maximized likelihood function, apart from a constant, is

$$L_{\max}^{-2/T} = \left|S_{00}\right| \prod_{i=1}^{r-s}\left(1-\tilde{\lambda}_i\right) \prod_{i=1}^{s}\left(1-\tilde{\rho}_i\right). \tag{33}$$

The proof of Theorem 1 is in the Appendix.

Theorem 2. *The likelihood ratio test statistic of the hypothesis* $H_0 : \beta = \left[H, \theta\right]$, $\alpha = A\psi$ *verses* $H\left(r\right)$ *is expressed as:*

$$LR\left(H_0 \mid H\left(r\right)\right) = T\left\{\sum_{i=1}^{r-s} \ln\left(1-\tilde{\lambda}_i\right) + \sum_{i=1}^{s} \ln\left(1-\tilde{\rho}_i\right) - \sum_{j=1}^{r} \ln\left(1-\hat{\lambda}_j\right)\right\}, \tag{34}$$

where $\left\{\hat{\lambda}_i\right\}_{i=1,r}$ *are from the unrestricted maximized likelihood in* (15)*, and is asymptotically distributed as* χ^2 *with r(p-m)+s(p-r) degrees of freedom.*

The proof of Theorem 2 is in the Appendix.

(7) $H_0 : \beta = H\phi,\ \alpha = \left[A, \tau\right]$

where H $p\times s$, A $p\times m$ are known; ϕ $s\times r$ is unknown; and τ $p\times(r\text{-}m)$ is unknown where $\tau = \bar{A}_\perp \psi$ with $\bar{A}_\perp$ $p\times(p\text{-}m)$ known and $\psi(p\text{-}m)\times(r\text{-}m)$ unknown, $m \le r \le s < p$.

This test combines Johansen's tests (1) and (4), that is, it tests the restriction that the cointegrating vectors share *p-s* linear restrictions and *m* of the adjustment vectors are assumed known (with the remaining *r-m* orthogonal to them). This test would be used, for example, to determine if some variable in the system did not enter any of the cointegrating relationships or if two variables entered the cointegrating relationships as the spread between them, and to test simultaneously that some of the cointegrating vectors only appear in the equation for one of the variables.

The first step in calculating the test statistic and restricted coefficient estimates is to split the reduced rank regression into variation independent parts

$$\begin{aligned}\bar{A}' R_{0t} &= \varphi_1' H' R_{1t} + \bar{A}' \varepsilon_t \\ A_\perp' R_{0t} &= \psi\varphi_2' H' R_{1t} + A_\perp' \varepsilon_t\end{aligned}, \tag{35}$$

where ϕ is partitioned conformably with α as $\left[\phi_1, \phi_2\right]$. In order to derive the test statistics and to estimate the restricted parameters under this hypothesis it is again necessary to define a new set of residual vectors and transform the product moment matrices, S_{ij}. Fixing ϕ_2 and ϕ, define the residual vector

$$R_{kt} = A_\perp' R_{0t} - \psi\phi_2' H' R_{1t}. \tag{36}$$

One can then define the notation $S_{1k} = \frac{1}{T}\sum_{t=1}^{T} R_{1t}R'_{kt}$ and so on, and define the set of product moment matrices:

$$S_{ij.k} = S_{ij} - S_{ik}S_{kk}^{-1}S_{kj},\ i, j = 0,1. \tag{37}$$

The restricted estimators and the likelihood ratio test statistic and its asymptotic distribution are summarized in the following theorems.

Theorem 3. *Under the hypothesis* $H_0 : \beta = H\phi,\ \alpha = [A, \tau]$ *where* H *p×s,* A *p×m are known;* ϕ *s×r is unknown; and* τ *p×(r-m) is unknown where* $\tau = \overline{A}_{\perp}\psi$ *with* $\overline{A}_{\perp}$ *p×(p-m) known and* ψ *(p-m)×(r-m) unknown,* $m \le r \le s < p$*; the maximum likelihood estimators are found by the following steps:*

Solve the eigenvalue problem

$$\left|\lambda H'S_{11}H - H'S_{10}A_{\perp}\left(A'_{\perp}S_{00}A_{\perp}\right)^{-1}A'_{\perp}S_{01}H\right| = 0 \tag{38}$$

for eigenvalues $1 \ge \tilde{\lambda}_1 \ge \ldots \ge \tilde{\lambda}_s \ge 0$ *and corresponding eigenvectors* $\tilde{V} = (\tilde{v}_1, \ldots, \tilde{v}_s)$*, normalized so that* $\tilde{V}'H'S_{11}H\tilde{V} = I_s$*; and solve the eigenvalue problem*

$$\left|\rho H'S_{11.k}H - H'S_{10.k}A\left(A'S_{00.k}A\right)^{-1}A'S_{01.k}\right| = 0 \tag{39}$$

for eigenvalues $1 \ge \tilde{\rho}_1 \ge \ldots \ge \tilde{\rho}_m > \tilde{\rho}_{m+1} = \ldots = \tilde{\rho}_s = 0$.

Then the restricted estimators are

$$\hat{\phi}_2 = (\tilde{v}_1, \ldots, \tilde{v}_{r-m}) \tag{40}$$

$$\hat{\beta}_2 = H\hat{\phi}_2 \tag{41}$$

$$\hat{\psi} = A'_{\perp}S_{01}H\hat{\phi}_2 \tag{42}$$

$$\hat{\phi}_1 = \left(H'S_{11.k}H\right)^{-1}H'S_{10.k}\overline{A} \tag{43}$$

$$\hat{\beta} = \left[\hat{\beta}_1, \hat{\beta}_2\right] = \left[H\left(H'S_{11.k}H\right)^{-1}H'S_{10.k}\overline{A}, H\hat{\phi}_2\right] \tag{44}$$

$$\hat{\alpha} = [A, \hat{\tau}] = \left[A, \overline{A}_{\perp}\hat{\psi}\right] = \left[A, A_{\perp}\left(A'_{\perp}A_{\perp}\right)^{-1}A'_{\perp}S_{01}\hat{\beta}_2\right], \tag{45}$$

where $S_{ij.k} = S_{ij} - S_{ik}S_{kk}^{-1}S_{kj},\ i, j = 0,1$ *is calculated from* (37) *evaluated at* $\hat{\phi}_2, \hat{\psi}$.

The maximized likelihood function, apart from a constant, is

$$L_{\max}^{-2/T} = \frac{\left|A'S_{00.k}A\right|\left|A'_{\perp}S_{00}A_{\perp}\right|}{\left|A'A\right|\left|A'_{\perp}A_{\perp}\right|}\prod_{i=1}^{r-m}\left(1-\tilde{\lambda}_i\right)\prod_{i=1}^{m}\left(1-\tilde{\rho}_i\right). \tag{46}$$

The proof of Theorem 3 is in the Appendix.

Theorem 4. *The likelihood ratio test statistic of the hypothesis* $H_0 : \beta = H\phi,\ \alpha = [A, \tau]$ *verses* $H(r)$ *is expressed as:*

$$LR\left(H_0 \mid H(r)\right) = T\left\{\ln\left[\frac{\left|A'S_{00.k}A\right|\left|A_\perp'S_{00}A_\perp\right|}{\left|A'A\right|\left|A_\perp'A_\perp\right|}\right] - \ldots \right. \\ \left. \ln\left|S_{00}\right| + \sum_{i=1}^{r-m}\ln\left(1-\tilde{\lambda}_i\right) + \sum_{i=1}^{m}\ln\left(1-\tilde{\rho}_i\right) - \sum_{j=1}^{r}\ln\left(1-\hat{\lambda}_j\right)\right\}, \tag{47}$$

where $\left\{\hat{\lambda}_i\right\}_{i=1,r}$ *are from the unrestricted maximized likelihood in* (15)*, and is asymptotically distributed as* χ^2 *with m(p-r)+r(p-s)degrees of freedom.*

The proof of Theorem 4 is in the Appendix.

Next, a hypothesis test on $\Pi = \alpha\beta'$ of the form $\Pi = \Pi_1 + \Pi_2$ is presented in which $\Pi_1 = AH'$ is known. This test, which combines tests (2) and (4), implies one is testing that both a subset of the cointegrating vectors and the associated adjustment vectors are known. It might seem too optimistic or restrictive to believe one might not only know certain cointegrating vectors but also know the adjustments to them. A test of this sort, however, might be useful as the end of a general-to-simple strategy for testing structural hypotheses or for testing very specific theoretical implications. More usefully, one might estimate the cointegrating relationships and adjustment vectors from a subset of a system of variables and then desire to test whether these estimated relationships hold in the full system of variables.

(8) $H_0: \beta = [H, \theta],\ \alpha = [A, \tau]$

where both H, A are known $p \times s$ matrices with $s<r$, and the unknown parameter matrices are orthogonal to H, A: $\theta = \bar{H}_\perp \phi$, $\tau = \bar{A}_\perp \psi$ with $\bar{H}_\perp$, $\bar{A}_\perp$ $p \times (p-s)$ known and ϕ, ψ $(p-s) \times (r-s)$ unknown. This implies $\Pi = AH' + \tau\theta' = AH' + \bar{A}_\perp \psi \phi' \bar{H}_\perp'$.

Define the vector of residuals

$$R_{kt} = R_{0t} - AH'R_{1t}. \tag{48}$$

The reduced rank regression (16) is split into

$$\begin{aligned} A'R_{kt} &= A'\varepsilon_t \\ A_\perp'R_{kt} &= \psi_1 \phi' \bar{H}_\perp' R_{1t} + A'\varepsilon_t \end{aligned}. \tag{49}$$

In order to derive the test statistics and to estimate the restricted parameters under this hypothesis it is again necessary to define a new set of residual vectors and transform the product moment matrices, $S_{ik} = \frac{1}{T}\sum_{t=1}^{T} R_{it}R_{kt}'$, $i = 1, k$ and so on, and also define the product moment matrices, $S_{ij.A} = S_{ij} - S_{ik}A\left(A'S_{kk}A\right)^{-1}A'S_{kj}$, $i = 1, k$.

The restricted estimators and the likelihood ratio test statistic and its asymptotic distribution are summarized in the following theorem.

Theorem 5. *Under the hypothesis* $H_0: \beta = [H, \theta],\ \alpha = [A, \tau]$ *where* H, A *are known* $p \times s$ *matrices;* θ *and* τ *are unknown* $p \times (r-s)$ *matrices such that* $\theta = \bar{H}_\perp \phi$ *and* $\tau = \bar{A}_\perp \psi$ *with* $\bar{H}_\perp$

and $\bar{A}_\perp$ *□p×(p-s) known and* ϕ, ψ *(p-s)×(r-s) unknown; the maximum likelihood estimators are found by the following steps:*

Solve the eigenvalue problem

$$\left|\lambda \bar{H}_\perp' S_{11.A} \bar{H}_\perp - \bar{H}_\perp' S_{1k.A} A_\perp \left(A_\perp S_{kk.A} A_\perp\right)^{-1} A_\perp' S_{k1.A} \bar{H}_\perp\right| = 0 \tag{50}$$

for eigenvalues $1 \geq \tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_{s-r} \geq \tilde{\lambda}_{s-r+1} = \ldots = \tilde{\lambda}_{p-s} = 0$ *and corresponding eigenvectors* $\tilde{V} = \left(\tilde{v}_1, \ldots, \tilde{v}_{p-s}\right)$, *normalized so that* $\tilde{V}' \bar{H}_\perp' S_{11.A} \bar{H}_\perp \tilde{V} = I_{p-r}$.

Then the restricted estimators are

$$\hat{\phi} = \left(\tilde{v}_1, \ldots, \tilde{v}_{r-s}\right) \tag{51}$$

$$\beta = \left[\hat{\beta}_1, \hat{\beta}_2\right] = \left[H, \hat{\theta}\right] = \left[H, \bar{H}_\perp \hat{\phi}\right] \tag{52}$$

$$\hat{\psi} = A_\perp' S_{k1.A} \bar{H}_\perp \hat{\phi}, \tag{53}$$

$$\alpha = \left[A, \tau\right] = \left[A, \bar{A}_\perp \psi\right] = \left[A, A_\perp \left(A_\perp' A_\perp\right)^{-1} A_\perp' S_{k1.A} \hat{\beta}_2\right] \tag{54}$$

and the maximized likelihood function, apart from a constant, is

$$L_{\max}^{-2/T} = \left|S_{kk}\right| \prod_{i=1}^{r-s} \left(1 - \tilde{\lambda}_i\right). \tag{55}$$

The proof of Theorem 5 is in the Appendix.

Theorem 6. *The likelihood ratio test statistic of the hypothesis* $H_0 : \beta = \left[H, \theta\right]$, $\alpha = \left[A, \tau\right]$ *verses* $H(r)$ *is expressed as:*

$$LR\left(H_0 \mid H(r)\right) = T\left\{\ln\left|S_{kk}\right| - \ln\left|S_{00}\right| + \sum_{i=1}^{r-s} \ln\left(1 - \tilde{\lambda}_i\right) - \sum_{j=1}^{r} \ln\left(1 - \hat{\lambda}_j\right)\right\}, \tag{56}$$

where $\left\{\hat{\lambda}_i\right\}_{i=1,r}$ *are from the unrestricted maximized likelihood in* (15), *and is asymptotically distributed as* χ^2 *with* $2ps\text{-}s^2$ *degrees of freedom.*

The proof of Theorem 6 is in the Appendix.

4. Testing Restrictions on $\alpha_\perp$ and $\beta_\perp$

Separating an economic time series into permanent (long run) components and cyclical (short run, temporary, transitory) components has been used in many contexts in economics. Methods proposed include decomposing the series into a deterministic trend component and a stationary cyclical component. Muth [33] uses the long-run forecast of a geometric distributed lag, that is, the permanent component is the long-run forecast after the dynamics (modeled as a distributed lag) have run their course. Beveridge and Nelson [1] use the Wold [47] decomposition to generalize this to ARIMA models, defining the permanent component to be a multiple of the random walk component of the series. This method, too, implies that the permanent component of the series in period *t* is the long-run forecast of the time series made in period *t*. Watson [44] uses unobserved components ARIMA models based on Watson and Engle's [46] methods. Quah [37] develops a permanent-transitory (P-T)

decomposition to derive lower bounds for the relative size of the permanent component of a series and showed that restricting it to be a random walk maximizes the size of the lower bound.

Sims [40] introduced vector autoregressions to empirical economics as a flexible multivariate dynamic framework to which the Beveridge-Nelson [1] decomposition can be extended (see [42]). In cointegrated systems, several methods have been proposed to decompose the individual time series into their permanent and cyclical components. The importance of multivariate information sets for this sort of analysis is argued in Cochrane [3]. Stock and Watson [42], Johansen [20], and Granger and Gonzalo [12] split a system of p cointegrated time series into p-r common stochastic trends (where r is the number of cointegrating relationships), linear combinations of which form the permanent components of the individual time series. The cyclical components are some combination of the cointegrating relationships, plus, if the common stochastic trends are assumed to be random walks, other stationary components. See [36] for a discussion of the relationship among these definitions and with the notion of common features by Vahid and Engle [43] and Engle and Kozicki [8].

The orthogonal complements of β and α are used to construct the common stochastic trends and the permanent components of a cointegrated model. Kasa [27] proposes $\beta_\perp' X_t$ as the p-r common stochastic trends and $\beta_\perp \left(\beta_\perp' \beta_\perp\right)^{-1} \beta_\perp' X_t$ as the permanent components of the individual variables in the system. Gonzalo and Granger [12] propose $\alpha_\perp' X_t$ as the common stochastic trends in the system and $\beta_\perp \left(\alpha_\perp' \beta_\perp\right)^{-1} \alpha_\perp' X_t$ as the permanent components; Johansen [22] proposes the random walks $\alpha_\perp' \Gamma(L) X_t$ as the common stochastic trends and random walks $\beta_\perp \left(\alpha_\perp' \Gamma(1) \beta_\perp\right)^{-1} \alpha_\perp' \Gamma(L) X_t$ as the permanent components.

There is no econometric reason why one definition of a common stochastic trend and permanent component is necessarily any better than another; one needs economic justifications to choose among them. One interpretation of the cointegrating relationships, β, derived from Johansen's methodology is that they are the r maximally canonically correlated linear combinations of ΔX_t and X_{t-1}. So, Kasa's common stochastic trends would be the p-r minimally canonically correlated linear combinations; there, however, is no strong economic justification for choosing these linear combinations as the common stochastic trends. The Gonzalo and Granger formulation has the advantage that the cointegrating relationships and transitory components have no long-run effect on the common stochastic trends and permanent components. In the Johansen version, the common stochastic trends and permanent components are random walks (like the univariate Beveridge-Nelson decomposition), and the permanent components of the variables can be seen as the long-run forecasts of the variables once the dynamics have worked out themselves. In the Johansen definition, however, unlike the Gonzalo and Granger method, the cointegrating relationships and transitory components can have a permanent effect on the common stochastic trends and the permanent components.

Recall that β and α are $p \times r$ matrices of full column rank, that is, the columns of β and α lie in r-dimensional subspaces of $\mathbb{R}^p$. The likelihood ratio tests in section 3 for restrictions on the cointegrating vectors and on their disequilibrium adjustment vectors were of two general types: The first imposes linear relationships on all the vectors, and the second assumes that a subset of the vectors are known. Johansen [20] shows that since one actually

estimates the space spanned by the cointegrating vectors, $sp(\beta)$, restrictions on cointegrating vectors are restrictions on the space they span. The restriction that the r vectors in β share p-s linear restrictions, that is, $\beta = H\phi$ where H is a known $p\times s$ matrix of full column rank and ϕ is an unknown $s\times r$ matrix, can be represented geometrically as $sp(\beta) \subset sp(H)$. This implies the columns of β are restricted to lie in a given s-dimensional subspace of $\mathbb{R}^p$ [19]. The restriction that m of the cointegrating relationships are known, that is, $\beta = [h, \phi]$ where h contains the known $p\times m$ relationships and $\phi = h_\perp \vartheta$ $p\times(r\text{-}m)$ is unknown, can be represented geometrically as $sp(h) \subset sp(\beta)$ [20]. This implies that the known vectors lie in an m-dimensional subspace of the space spanned by the vectors in β. These two restrictions can be written $sp(h) \subset sp(\beta) \subset sp(H)$.

Restrictions placed on cointegrating vectors or on their adjustment vectors imply that restrictions are imposed on the space spanned by their orthogonal complements as well [20]. The restriction that $sp(\beta) \subset sp(H)$ implies $sp(H_\perp) \subset sp(\beta_\perp)$, where the orthogonal complements $\beta_\perp$ and $H_\perp$ are $p\times(p\text{-}r)$ and $p\times(p\text{-}s)$ matrices, respectively, of full column rank. This means that a subset of p-s of the p-r vectors in $\beta_\perp$ are known, namely those contained in $H_\perp$. Thus, the test $\beta = H\phi$ implies a test on its orthogonal complement of the form $\beta_\perp = [H_\perp, \theta]$ for which θ is an unknown $p\times(s\text{-}r)$ matrix of rank s-r.

Similarly, $sp(h) \subset sp(\beta)$ implies $sp(\beta_\perp) \subset sp(h_\perp)$, where $h_\perp$ is a $p\times(p\text{-}m)$ matrix of full column rank; that is, the vectors in $\beta_\perp$ share the $(p\text{-}m)$ linear restrictions implied by $h_\perp$. Thus, a test of the form $\beta = [h, \phi]$ implies a test on its orthogonal complement of the form $\beta_\perp = h_\perp \theta$ for which θ is an unknown $(p\text{-}m)\times(p\text{-}r)$ matrix of rank p-r.

With minor modifications to the tests in section 3, we may more explicitly state the implications for the orthogonal complements and reformulate them as tests on the orthogonal complements, that is, use the tests in section 3 as tests on the orthogonal complements.

Theorem 7. *For (1) $H_0 : \beta = H\phi$ where H $p\times s$ is known and ϕ $s\times r$ is unknown, $r\leq s<p$ one may choose*

$$\beta_\perp = [H_\perp, \bar{H}\phi_\perp] \tag{57}$$

where $\bar{H} \equiv H(H'H)^{-1}$. Further, one can test the hypothesis

$$\beta_\perp = [G, \bar{G}_\perp \theta] \tag{58}$$

where G $p\times q$ is known and θ $(p\text{-}q)\times(p\text{-}q\text{-}r)$ is unknown by transforming this problem into H_0 above setting $H = G_\perp$ and $s=p\text{-}q$. That is, one may test the hypothesis that certain $\beta_\perp$ are known and the remaining elements of $\beta_\perp$ are orthogonal to the known vectors.

To check that $\beta_\perp$ is indeed an orthogonal complement of β, one must verify that

$$\beta_\perp'\beta = 0_{(p-r)\times r}$$

$$\beta_\perp'\beta = [H_\perp, \bar{H}\phi_\perp]' H\phi = \begin{bmatrix} H_\perp' H\phi \\ \phi_\perp' \bar{H}' H\phi \end{bmatrix} = \begin{bmatrix} 0\phi \\ \phi_\perp' I_s \phi \end{bmatrix} = \begin{bmatrix} 0_{(p-s)\times r} \\ 0_{(s-r)\times r} \end{bmatrix} = 0_{(p-r)\times r}$$

Theorem 8. *For (2)* $H_0 : \beta = \left[H, \bar{H}_\perp \phi \right]$ *where H p×s is known and* ϕ *(p-s)×(r-s) is unknown, one may choose*

$$\beta_\perp = \left[H_\perp \phi_\perp \right]. \tag{59}$$

Thus, we can test the hypothesis

$$\beta_\perp = G\theta, \tag{60}$$

where G is a known p×q matrix and θ *is an unknown q×(p-r) matrix by transforming this problem into* H_0 *above setting* $H = G_\perp$ *and s=p-q. That is, one may test the hypothesis that the vectors in* $\beta_\perp$ *share the same p-s linear restrictions.*

Again, to check that $\beta_\perp$ is indeed an orthogonal complement of β, one must verify that $\beta_\perp' \beta = 0_{(p-r)\times r}$:

$$\beta_\perp' \beta = \phi_\perp' H_\perp' \left[H, \bar{H}_\perp \phi \right] = \left[\phi_\perp' H_\perp' H, \phi_\perp' H_\perp' \bar{H}_\perp \phi \right] = \left[\phi_\perp' 0, \phi_\perp' I_{(p-s)} \phi \right] = 0_{(p-r)\times r}.$$

Proofs of the above and following theorems are in the Appendix.

One can apply the ideas from the two examples above to tests (3) through (8) in section 3. The results are summarized below.

Theorem 9. *For (3)* $H_0 : \alpha = A\psi$ *where A p×m is known and* ψ *m×r is unknown, r≤m≤p, one may choose*

$$\alpha_\perp = \left[A_\perp, \bar{A}\psi_\perp \right], \tag{61}$$

where $\bar{A} \equiv A\left(A'A\right)^{-1}$. *Further, one can test the hypothesis*

$$\alpha_\perp = \left[B, \bar{B}_\perp \xi \right] \tag{62}$$

where B p×n is known and ξ *(p-n)×(p-n-r) is unknown by transforming this problem into* H_0 *above setting* $A = B_\perp$ *and m=p-n. That is, one may test the hypothesis that certain* $\alpha_\perp$ *are known and the remaining vectors in* $\alpha_\perp$ *are orthogonal to the known vectors.*

Theorem 10. *For (4)* $H_0 : \alpha = \left[A, \bar{A}_\perp \psi \right]$ *where A p×m is known and* ψ *(p-m)×r is unknown, one may choose*

$$\alpha_\perp = A_\perp \psi_\perp. \tag{63}$$

Further, one can test the hypothesis

$$\alpha_\perp = B\xi \tag{64}$$

where B is a known p×n matrix and ξ *is an unknown n×(p-r) matrix by transforming this problem into* H_0 *above setting* $A = B_\perp$ *and m=p-n. That is, one may test the hypothesis that the vectors in* $\alpha_\perp$ *share the same p-m linear restrictions.*

This test is equivalent to the hypothesis test H_{4b} in [12], which uses the dual of the eigenvalue problem for (4) used in Theorem 10.

The following four theorems allow one to combine restrictions on the orthogonal complements of the cointegrating vectors and of their disequilibrium adjustment vectors.

Theorem 11. *For (5)* $H_0 : \beta = H\phi,\ \alpha = A\psi$ *where H p×s, A p×m are known and* ϕ *s×r,* ψ *m×r are unknown, r≤s<p and r≤m<p, one may choose*

$$\beta_\perp = \left[H_\perp, \bar{H}\phi_\perp \right],\ \alpha_\perp = \left[A_\perp, \bar{A}\psi_\perp \right]. \tag{65}$$

Thus, we can test the hypothesis

$$\beta_\perp = \left[G, \bar{G}_\perp \theta \right],\ \alpha_\perp = \left[B, \bar{B}_\perp \xi \right] \tag{66}$$

where G p×q and B p×n are known matrices and θ *(p-q)×(p-q-r) and* ξ *(p-n)×(p-n-r) are unknown matrices by transforming this problem into* H_0 *above setting* $H = G_\perp$, $A = B_\perp$, *s=p-q, and m=p-n. This test allows one simultaneously to test for known* $\beta_\perp$ *vectors and for known* $\alpha_\perp$ *vectors.*

Theorem 12. *For (6)* $H_0 : \beta = H\phi,\ \alpha = \left[A, \bar{A}_\perp \psi \right]$ *where H p×s, A p×m are known and* ϕ *s×r,* ψ *(p-m)×(m-r) are unknown, m≤r≤s<p, one may choose*

$$\beta_\perp = \left[H_\perp, \bar{H}\phi_\perp \right],\ \alpha_\perp = A_\perp \psi_\perp. \tag{67}$$

Thus, we can test the hypothesis

$$\beta_\perp = \left[G, \bar{G}_\perp \theta \right],\ \alpha_\perp = B\xi \tag{68}$$

where G p×q and B p×n are known matrices and θ *(p-q)×(p-q-r) and* ξ *n×(p-r) are unknown matrices by transforming this problem into* H_0 *above setting* $H = G_\perp$, $A = B_\perp$, *s=p-q, and m=p-n. This test allows one simultaneously to test for known* $\beta_\perp$ *vectors and to place common linear restrictions on* $\alpha_\perp$.

Theorem 13. *For (7)* $H_0 : \beta = \left[H, \bar{H}_\perp \phi \right],\ \alpha = A\psi$ *where H p×s, A p×m are known and* ϕ *(p-s)×(r-s),* ψ *m×r are unknown, s≤r≤m<p, one may choose*

$$\beta_\perp = \left[H_\perp \phi_\perp \right],\ \alpha_\perp = \left[A_\perp, \bar{A}\psi_\perp \right] \tag{69}$$

Thus, we can test the hypothesis

$$\beta_\perp = G\theta,\ \alpha_\perp = \left[B, \bar{B}_\perp \xi \right] \tag{70}$$

where G p×q and B p×n are known matrices and θ *q×(p-r) and* ξ *(p-n)×(p-n-r) are unknown matrices, by transforming this problem into* H_0 *above setting* $H = G_\perp$, $A = B_\perp$, *s=p-q, and m=p-n. This test allows one simultaneously to test for common linear restrictions on the* $\beta_\perp$ *vectors and for known* $\alpha_\perp$ *vectors.*

Theorem 14. *For (8)* $H_0 : \beta = \left[H, \bar{H}_\perp \phi \right],\ \alpha = \left[A, \bar{A}_\perp \psi \right]$ *where H p×s, A p×s are known and* ϕ, ψ *(p-s)×(r-s) are unknown, s≤r<p, one may choose*

$$\beta_{\perp} = \left[H_{\perp} \phi_{\perp} \right],\ \alpha_{\perp} = A_{\perp} \psi_{\perp} \tag{71}$$

Thus, we can test the hypothesis

$$\beta_{\perp} = G\theta,\ \alpha_{\perp} = B\xi \tag{72}$$

where G p×q and B p×n are known matrices and θ,ξ q×(p-r) are unknown matrices, by transforming this problem into H_0 *above setting* $H = G_{\perp}$, $A = B_{\perp}$, *s=p-q. This test allows one simultaneously to test for common linear restrictions on the* $\beta_{\perp}$ *vectors and to place common linear restrictions on the* $\alpha_{\perp}$ *vectors.*

These theorems allow one, in addition, to combine the tests on the cointegrating vectors and adjustment vectors with those on the respective orthogonal complements. For example, one could use Theorem 13 to test that the cointegrating vectors share certain linear restrictions (say, ratios or spreads, or that some subset of variables do not enter the cointegrating relationships) and that some subset of the common stochastic trends are known: $H_0 : \beta = H\phi,\ \alpha_{\perp} = \left[B, \bar{B}_{\perp} \xi \right]$. The tests (1) through (8) can be recast as tests of the hypotheses that are displayed below:

Test (1)	$\beta = H\phi$	$\beta_{\perp} = \left[G, \bar{G}_{\perp} \theta \right]$		
Test (2)	$\beta = \left[H, \bar{H}_{\perp} \phi \right]$	$\beta_{\perp} = G\theta$		
Test (3)	$\alpha = A\psi$	$\alpha_{\perp} = \left[B, \bar{B}_{\perp} \xi \right]$		
Test (4)	$\alpha = \left[A, \bar{A}_{\perp} \psi \right]$	$\alpha_{\perp} = B\xi$		
Test (5)	$\beta = H\phi$ $\alpha = A\psi$	$\beta = H\phi$ $\alpha_{\perp} = \left[B, \bar{B}_{\perp} \xi \right]$	$\beta_{\perp} = \left[G, \bar{G}_{\perp} \theta \right]$ $\alpha = A\psi$	$\beta_{\perp} = \left[G, \bar{G}_{\perp} \theta \right]$ $\alpha_{\perp} = \left[B, \bar{B}_{\perp} \xi \right]$
Test (6)	$\beta = \left[H, \bar{H}_{\perp} \phi \right]$ $\alpha = A\psi$	$\beta = \left[H, \bar{H}_{\perp} \phi \right]$ $\alpha_{\perp} = \left[B, \bar{B}_{\perp} \xi \right]$	$\beta_{\perp} = G\theta$ $\alpha = A\psi$	$\beta_{\perp} = G\theta$ $\alpha_{\perp} = \left[B, \bar{B}_{\perp} \xi \right]$
Test (7)	$\beta = H\phi$ $\alpha = \left[A, \bar{A}_{\perp} \psi \right]$	$\beta = H\phi$ $\alpha_{\perp} = B\xi$	$\beta_{\perp} = \left[G, \bar{G}_{\perp} \theta \right]$ $\alpha = \left[A, \bar{A}_{\perp} \psi \right]$	$\beta_{\perp} = \left[G, \bar{G}_{\perp} \theta \right]$ $\alpha_{\perp} = B\xi$
Test (8)	$\beta = \left[H, \bar{H}_{\perp} \phi \right]$ $\alpha = \left[A, \bar{A}_{\perp} \psi \right]$	$\beta = \left[H, \bar{H}_{\perp} \phi \right]$ $\alpha_{\perp} = B\xi$	$\beta_{\perp} = G\theta$ $\alpha = \left[A, \bar{A}_{\perp} \psi \right]$	$\beta_{\perp} = G\theta$ $\alpha_{\perp} = B\xi$

where $G = H_{\perp}, B = A_{\perp}, \theta = \phi_{\perp}, \zeta = \psi_{\perp}$, and vice versa.

5. Conclusion

This paper has two aims. The first is to develop three new hypothesis tests for combining structural hypotheses on cointegrating relationships and on their disequilibrium adjustment vectors in Johansen's [19] multivariate maximum likelihood cointegration framework. These tests possess closed-form solutions for parameter estimates under the null

hypothesis. The second is to demonstrate the implications that the tests for restrictions on the cointegration vectors and disequilibrium adjustment vectors have for the orthogonal complements of these quantities, and how these tests can be formulated as tests on the orthogonal complements. This is useful since the various specifications of multivariate common stochastic trends and permanent components are derived from these orthogonal complements. Thus, one may combine tests for restrictions on the long-run relationships represented by cointegrating relationships, the adjustments to them, and the common stochastic trends of a system of variables.

ACKNOWLEDGEMENT The genesis of this paper occurred, in part, during conversations with Professor Granger in his office.

References

[1] S. Beveridge and C.R. Nelson. A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'. *Journal of Monetary Economics*, 7:151-174, 1981.

[2] Y-W Cheung and K.S. Lai. Finite-Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration. *Oxford Bulletin of Economics and Statistics*, 55:313-328, 1993.

[3] J Cochrane. Univariate vs. Multivariate Forecasts of GNP Growth and Stock Returns: Evidence and Implications for the Persistence of Shocks, Detrending Methods, and Tests of the Permanent Income Hypothesis. Working Paper 3427, National Bureau of Economic Research, Cambridge, MA, 1994.

[4] J Cochrane. Permanent and Transitory Components of GNP and Stock Prices. *Quarterly Journal of Economics*, 241-265, 1994.

[5] J Cochrane. Shocks. *Carnegie-Rochester Conference Series on Public Policy*, 41:296-364, 1994.

[6] J Doornik. Testing General Restrictions on the Cointegrating Space. Working Paper, Nuffield College, 1995.

[7] R Engle and C Granger. Cointegration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55:251-276, 1987.

[8] R Engle and S Kozicki. Testing for Common Features. *Journal of Business and Economic Statistics*, 11(4):369-380, 1993.

[9] R Engle and B Yoo. Forecasting and Testing in Cointegrated Systems. *Journal of Econometrics*, 35:143-159, 1987.

[10] R Engle and B Yoo. Cointegrated Economic Time Series: An Overview with New Results. In R Engle and C Granger, editors, *Long-Run Economic Relations, Readings in Cointegration.* Oxford University Press, New York, 1987.

[11] J Gonzalo. Five Alternative Methods of Estimating Long-Run Equilibrium Relationships. *Journal of Econometrics*, 60:203-233, 1994.

[12] J Gonzalo and C Granger. Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business and Economic Statistics*, 13:27-35, 1995.

[13] J Gonzalo and T Lee. Pitfalls in Testing for Long Run Relationships. *Journal of Econometrics*, 86(1):129-154, 1998.

[14] C Granger. Some Properties of Time Series Data and Their Use in Econometric Model Specification. *Journal of Econometrics*, 16:121-130, 1981.

[15] C Granger. Cointegrated Variables and Error Correction Models. Discussion Paper 83:13a, University of California, San Diego, 1983.

[16] C Granger and T Lee. Multicointegration. *Advances in Econometrics*, 8:71-84, 1990.

[17] J Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, NJ, 1994.

[18] A Haug. Testing Linear Restrictions on Cointegrating Vectors: Sizes and Powers of Wald tests in Finite Samples. *Econometric Theory*, 18:505–524, 2002.

[19] S Johansen. Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, 12:231-254, 1988.

[20] S Johansen. Likelihood Based Inference on Cointegration: Theory and Applications. unpublished lecture notes, University of Copenhagen, Institute of Mathematical Statistics, 1989.

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

[22] S Johansen. Identifying Restrictions of Linear Equations with Applications to Simultaneous Equations and Cointegration. *Journal of Econometrics*, 1995, 69:111-132, 1995.

[23] S Johansen. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford, 1996.

[24] S Johansen. A Bartlett Correction Factor for Tests on the Cointegrating Relations. *Econometric Theory*, 16:740-778, 2000.

[25] S Johansen and K Juselius. Maximum Likelihood Estimation and Inference on Cointegration: with Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics*, 52:169-210, 1990.

[26] S Johansen and K Juselius. Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and UIP for UK. *Journal of Econometrics*, 53:211-244, 1992.

[27] K Kasa. Common Stochastic Trends in International Stock Markets. *Journal of Monetary Economics*, 29:95-124, 1992.

[28] R King, C Plosser, and S Robelo. Production, Growth, and Business Cycles: II. New Directions. *Journal of Monetary Economics*, 21:309-341, 1988.

[29] R King, C Plosser, J Stock, and M Watson. Stochastic Trends and Economic Fluctuations. *The American Economic Review*, 81:819-840, 1991.

[30] T Konishi and C Granger. Separation in Cointegrated Systems. Discussion Paper 92-51, University of California, San Diego, 1992.

[31] J MacKinnon, A Haug, and L Michelis. Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. *Journal of Applied Econometrics*, 14(5):563-577, 1999.

[32] R Mosconi and C Giannini. Non-causality in Cointegrated Systems: Representation, Estimation, and Testing. *Oxford Bulletin of Economics and Statistics*, 54:399-417, 1992.

[33] J Muth. Optimal Properties of Exponentially Weighted Forecasts. *Journal of the American Statistical Association*, 55:299-306, 1960.

[34] M Osterwald-Lenum. A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics*, 54:461-472, 1992.

[35] P Phillips. Optimal Inference in Cointegrated Systems. *Econometrica*, 56:1021-1044, 1991.

[36] T Proietti. Short Run Dynamics in Cointegrated Systems. *Oxford Bulletin of Economics and Statistics*, 59(3):405-422, 1997.

[37] D Quah. The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds. *Econometrica*, 60:107-118, 1992.

[38] H Reimers. Comparisons of Tests for Multivariate Cointegration. *Statistical Papers*, 33:335-359, 1992.

[39] G Reinsel and S Ahn. Vector Autoregressive Models with Unit Roots and Reduced Rank Structure: Estimation, Likelihood Ratio Tests, and Forecasting. *Journal of Time Series Analysis*, 13:353-375, 1990.

[40] C Sims. Macroeconomics and Reality. *Econometrica*, 48:1-48, 1980.

[41] R Solow. *Growth Theory: An Exposition*. Oxford, Clarendon Press, 1970.

[42] J Stock and M Watson. Testing for Common Trends. *Journal of the American Statistical Association*, 83:1097-1107, 1988.

[43] F Vahid and R Engle. Common Trends and Common Cycles. *Journal of Applied Econometrics*, 8(4):341-360, 1993.

[44] M Watson. Univariate Detrending Methods with Stochastic Trends. *Journal of Monetary Economics*, 18:49-75, 1986.

[45] M Watson. Vector Autoregressions and Cointegration. In R Engle and D. McFadden, editors, *The Handbook of Econometrics, 4*. Elsevier, Amsterdam, The Netherlands, 1995.

[46] M Watson and R Engle. Alternative algorithms for the Estimation of Dynamic Factor, MIMIC, and Varying Coefficient Regression Models. *Journal of Econometrics*, 23:485-500, 1983.

[47] H Wold. *A Study in the Analysis of Stationary Time Series*. Almqvist and Wiksell, Uppsala, Sweden, 1938.

Appendix

Proof of Theorem 1. $H_0 : \beta = \left[H, \bar{H}_\perp \phi \right]$, $\alpha = A\psi$ where H $p \times s$, A $p \times m$ are known and ϕ $(p-s) \times (r-s)$, ψ $m \times r$ are unknown, $r \leq m < p$.

The reduced rank regression from (6) is

$$R_{0t} = A\psi_1 H' R_{1t} + A\psi_2 \phi' \bar{H}_\perp' R_{1t} + \hat{\varepsilon}_t, \tag{73}$$

where ψ is partitioned conformably with β as $\left[\psi_1, \psi_2\right]$, and is split into

$$\bar{A}' R_{0t} = \psi_1 H' R_{1t} + \psi_2 \phi' \bar{H}_\perp' R_{1t} + \bar{A}' \hat{\varepsilon}_t \tag{74}$$

and

$$A_\perp' R_{0t} = A_\perp' \hat{\varepsilon}_t . \tag{75}$$

This allows one to factor the likelihood function into a marginal part based on (75) and a factor based on (74) conditional on (75):

$$\bar{A}'R_{0t} = \psi_1 H'R_{1t} + \psi_2 \phi' \bar{H}'_{\perp} R_{1t} + \omega A'_{\perp} R_{0t} + \bar{A}'\hat{\varepsilon}_t - \omega A'_{\perp}\hat{\varepsilon}_t \tag{76}$$

where

$$\omega = \Omega_{AA_{\perp}} \Omega^{-1}_{A_{\perp}A_{\perp}} = \bar{A}'\Omega A_{\perp} \left(A'_{\perp}\Omega A_{\perp} \right)^{-1}. \tag{77}$$

The parameters in the two equations are variation independent with independent errors, and the maximized likelihood will be the product of the maxima of the two factors.

The maximum of the likelihood function for the factor corresponding to the marginal distribution of $A'_{\perp}R_{0t}$ is, apart from a constant, $L^{-2/T}_{\max M} = \dfrac{\left|\hat{\Omega}_{A_{\perp}A_{\perp}}\right|}{\left|A'_{\perp}A_{\perp}\right|}$. The denominator is estimated by $\hat{\Omega}_{A_{\perp}A_{\perp}} = \left(A'_{\perp}\hat{\Omega}A_{\perp} \right) = A'_{\perp}S_{00}A_{\perp}$ thus,

$$L^{-2/T}_{\max M} = \frac{\left|A'_{\perp}S_{00}A_{\perp}\right|}{\left|A'_{\perp}A_{\perp}\right|} \tag{78}$$

Analysis of the factor of the likelihood function that corresponds to the distribution of $\bar{A}'R_{0t}$ conditional on $A'_{\perp}R_{0t}$ and R_{1t} is found by reduced rank regression. It is equivalent to maximizing the concentrated conditional factor as function of the unknown parameter matrix ϕ. First, one estimates ω by fixing ψ_1, ψ_2, and ϕ and regressing $\bar{A}'R_{0t} - \psi_1 H'R_{1t} - \psi_2 \phi' \bar{H}'_{\perp} R_{1t}$ on $A'_{\perp}R_{0t}$. This yields

$$\hat{\omega}(\psi_1, \psi_2, \phi) = \left(\bar{A}'S_{00}A_{\perp} - \psi_1 H'S_{10}A_{\perp} - \psi_2 \phi' \bar{H}'_{\perp} S_{10} A_{\perp} \right) \left(A'_{\perp}S_{00}A_{\perp} \right)^{-1}. \tag{79}$$

This allows one to correct for $A'_{\perp}R_{0t}$ in (75) by forming new residual vectors

$$R_{it.A_{\perp}} = R_{it} - S_{i0}A_{\perp} \left(A'_{\perp}S_{00}A_{\perp} \right)^{-1} A'_{\perp}R_{0t}, i = 0,1 \tag{80}$$

and product moment matrices

$$\begin{aligned} S_{ij.A_{\perp}} &= \frac{1}{T}\sum_{t=1}^{T} R_{it.A_{\perp}} R'_{jt.A_{\perp}} \\ &= S_{ij} - S_{i0}A_{\perp} \left(A'_{\perp}S_{00}A_{\perp} \right)^{-1} A'_{\perp}S_{0j}, i, j = 0,1 \end{aligned} \tag{81}$$

Thus we can write the conditional regression equation (76) as

$$\bar{A}'R_{0t.A_{\perp}} = \psi_1 H'R_{1t.A_{\perp}} + \psi_2 \phi' \bar{H}'_{\perp} R_{1t.A_{\perp}} + \bar{A}'\hat{\varepsilon}_t - \hat{\omega}A'_{\perp}\hat{\varepsilon}_t. \tag{82}$$

To successively concentrate the conditional likelihood until it is solely a function of ϕ, one fixes ψ_2 and ϕ and then estimates ψ_1 by regressing $\bar{A}'R_{0t.A_{\perp}} - \psi_2 \bar{H}'_{\perp} R_{1t.A_{\perp}}$ on $H'R_{1t.A_{\perp}}$ to get

$$\hat{\psi}_1 = \left(\bar{A}'S_{01.A_{\perp}}H - \psi_2 \phi' \bar{H}'_{\perp} S_{11.A_{\perp}} H \right) \left(H'S_{11.A_{\perp}}H \right)^{-1}. \tag{83}$$

One then corrects $R_{it.A_{\perp}}$ for $H'R_{1t.A_{\perp}}$ by forming new residuals

$$R_{it.A_{\perp}.H} = R_{it.A_{\perp}} - S_{i1.A_{\perp}}H \left(H'S_{11.A_{\perp}}H \right)^{-1} H'R_{1t.A_{\perp}}, \quad i = 0,1 \tag{84}$$

and product moment matrices

$$\begin{aligned} S_{ij.A_{\perp}.H} &= \frac{1}{T}\sum_{t=1}^{T} R_{it.A_{\perp}.H} R'_{jt.A_{\perp}.H} \\ &= S_{ij.A_{\perp}} - S_{i1.A_{\perp}}H \left(H'S_{11.A_{\perp}}H \right)^{-1} H'S_{1j.A_{\perp}}, \quad i, j = 0,1 \end{aligned} \tag{85}$$

Thus one can rewrite (82) as

$$\bar{A}'R_{0t.A_\perp.H} = \psi_2\phi'\bar{H}'_\perp R_{1t.A_\perp.H} + \hat{u}_t, \tag{86}$$

for which

$$\hat{u}_t = \bar{A}'\hat{\varepsilon}_t - \hat{\omega}A'_\perp\hat{\varepsilon}_t. \tag{87}$$

Fixing ϕ, one estimates ψ_2 by regressing $\bar{A}'R_{0t.A_\perp.H}$ on $\varphi'\bar{H}'_\perp R_{1t.A_\perp.H}$. This gives

$$\hat{\psi}_2 = \bar{A}'S_{01.A_\perp.H}\bar{H}_\perp\phi\left(\phi'\bar{H}'_\perp S_{11.A_\perp.H}\bar{H}_\perp\phi\right)^{-1} \tag{88}$$

and

$$\hat{u}_t = \bar{A}'R_{0t.A_\perp.H} - \bar{A}'S_{01.A_\perp.H}\bar{H}_\perp\phi\left(\phi'\bar{H}'_\perp S_{11.A_\perp.H}\bar{H}_\perp\varphi\right)^{-1}\phi'\bar{H}'_\perp R_{1t.A_\perp.H}. \tag{89}$$

The factor of the maximized likelihood corresponding to the conditional distribution is, apart from a constant,

$$L_{\max C}^{-2/T} = \frac{\left|\hat{\Omega}_{AA.A_\perp}\right|}{\left|\bar{A}'\bar{A}\right|} \tag{90}$$

where

$$\begin{aligned}\Omega_{AA.A_\perp} &= \Omega_{AA} - \Omega_{AA_\perp}\Omega^{-1}_{A_\perp A_\perp}\Omega_{A_\perp A} \\ &= \bar{A}'\Omega\bar{A}' - \bar{A}'\Omega A_\perp\left(A'_\perp\Omega A_\perp\right)^{-1}A'_\perp\Omega\bar{A}'\end{aligned}. \tag{91}$$

The maximum likelihood estimate of the conditional variance matrix is

$$\begin{aligned}\hat{\Omega}_{AA.A_\perp} &= \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t\hat{u}'_t \\ &= \bar{A}'S_{00.A_\perp.H}\bar{A} - \bar{A}'S_{01.A_\perp.H}\bar{H}_\perp\phi\left(\phi'\bar{H}'_\perp S_{11.A_\perp.H}\bar{H}_\perp\phi\right)^{-1}\phi'\bar{H}'_\perp S_{10.A_\perp.H}\bar{A}\end{aligned} \tag{92}$$

which gives the maximized conditional likelihood

$$L(\phi)^{-2/T}_{\max C} = \frac{\left|\bar{A}'S_{00.A_\perp.H}\bar{A} - \bar{A}'S_{01.A_\perp.H}\bar{H}_\perp\phi\left(\phi'\bar{H}'_\perp S_{11.A_\perp.H}\bar{H}_\perp\phi\right)^{-1}\phi'\bar{H}'_\perp S_{10.A_\perp.H}\bar{A}\right|}{\left|\bar{A}'\bar{A}\right|}. \tag{93}$$

The maximized likelihood function is the product between the maximized conditional factor and maximized marginal factor, for which the only unknown parameters are contained in φ; one then has (and noting that $\bar{A} \equiv A(A'A)^{-1}$ implies $\bar{A}'\bar{A} = A'A$)

$$L(\phi)^{-2/T}_{\max} = \frac{\left|A'_\perp S_{00}A_\perp\right|}{\left|A'_\perp A_\perp\right|\left|A'A\right|}\times\ldots \left|A'S_{00.A_\perp.H}A - A'S_{01.A_\perp.H}\bar{H}_\perp\phi\left(\phi'\bar{H}'_\perp S_{11.A_\perp.H}\bar{H}_\perp\phi\right)^{-1}\phi'\bar{H}'_\perp S_{10.A_\perp.H}A\right|. \tag{94}$$

From the matrix relationship for nonsingular *A* and *B*,

$$\begin{vmatrix}A & C\\ C' & B\end{vmatrix} = \left|A\right|\left|B - C'A^{-1}C\right| = \left|B\right|\left|A - CB^{-1}C'\right| \tag{95}$$

it follows that $\left|B - C'A^{-1}C\right| = \frac{|B|}{|A|}\left|A - CB^{-1}C'\right|$, and thus one can rewrite (94) as

$$L(\phi)^{-2/T}_{\max} = \frac{\left|A'_\perp S_{00} A_\perp\right|\left|A' S_{00.A_\perp.H} A\right|}{\left|A'_\perp A_\perp\right|\left|A'A\right|} \times \ldots$$
$$\frac{\left\|\phi' \bar{H}'_\perp S_{11.A_\perp.H} \bar{H}_\perp \phi - \phi' \bar{H}'_\perp S_{10.A_\perp.H} A\left(A' S_{00.A_\perp.H} A\right)^{-1} A' S_{01.A_\perp.H} \bar{H}_\perp \phi\right\|}{\left|\phi' \bar{H}'_\perp S_{11.A_\perp.H} \bar{H}_\perp \phi\right|}. \tag{96}$$

The variance-covariance matrix is then estimated by $\hat{\Omega} = \begin{bmatrix} A & \bar{A}_\perp \end{bmatrix} \begin{bmatrix} \hat{\Omega}_{AA} & \hat{\Omega}_{AA_\perp} \\ \hat{\Omega}_{A_\perp A} & \hat{\Omega}_{A_\perp A_\perp} \end{bmatrix} \begin{bmatrix} A & \bar{A}_\perp \end{bmatrix}'$, where the estimators of $\Omega_{A_\perp A_\perp}$, $\omega = \Omega_{AA_\perp}\Omega^{-1}_{A_\perp A_\perp}$, and $\Omega_{AA.A_\perp} = \Omega_{AA} - \Omega_{AA_\perp}\Omega^{-1}_{A_\perp A_\perp}\Omega_{A_\perp A}$ are used to recover $\hat{\Omega}_{A_\perp A_\perp}$, $\hat{\Omega}_{AA_\perp} = \hat{\omega}\hat{\Omega}_{A_\perp A_\perp}$, $\hat{\Omega}_{A_\perp A} = \hat{\Omega}'_{AA_\perp}$, and $\hat{\Omega}_{AA} = \hat{\Omega}_{AA.A_\perp} + \hat{\omega}\hat{\Omega}_{A_\perp A}$.
Maximizing the likelihood function is equivalent to minimizing the last factor of (96) with respect to ϕ. Following from Johansen and Juselius [25], here, one solves the eigenvalue problem

$$\left|\lambda \bar{H}'_\perp S_{11.A_\perp.H} \bar{H}_\perp - \bar{H}'_\perp S_{10.A_\perp.H} A\left(A' S_{00.A_\perp.H} A\right)^{-1} A' S_{01.A_\perp.H} \bar{H}_\perp\right| = 0 \tag{97}$$

for eigenvalues $1 \geq \tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_{p-s} \geq 0$ and eigenvectors $\tilde{V} = \left(\tilde{v}_1, \ldots, \tilde{v}_{p-s}\right)$ normalized so that $\tilde{V}'\bar{H}'_\perp S_{11.A_\perp.H}\bar{H}_\perp \tilde{V} = I_{p-s}$. Then $\hat{\phi} = \left(\tilde{v}_1, \ldots, \tilde{v}_{r-s}\right)$, from which one then can recover the parameters (29) to (32), and the maximized likelihood function, apart from a constant, is

$$L^{-2/T}_{\max} = \frac{\left|A'_\perp S_{00} A_\perp\right|\left|A' S_{00.A_\perp.H} A\right|}{\left|A'_\perp A_\perp\right|\left|A'A\right|} \prod_{i=1}^{r-s}\left(1 - \tilde{\lambda}_i\right). \tag{98}$$

Rewriting

$$\left|A' S_{00.A_\perp.H} A\right| = \left|A' S_{00.A_\perp} A - A' S_{01.A_\perp} H\left(H' S_{11.A_\perp} H\right)^{-1} H' S_{10.A_\perp} A\right|$$
$$= \frac{\left|A' S_{00.A_\perp} A\right|\left|H' S_{11.A_\perp} H - H' S_{10.A_\perp} A\left(A' S_{00.A_\perp} A\right)^{-1} A' S_{01.A_\perp} H\right|}{\left|H' S_{11.A_\perp} H\right|} \tag{99}$$

and noting

$$\frac{\left|H' S_{11.A_\perp} H - H' S_{10.A_\perp} A\left(A' S_{00.A_\perp} A\right)^{-1} A' S_{01.A_\perp} H\right|}{\left|H' S_{11.A_\perp} H\right|} = \prod_{i=1}^{s}\left(1 - \tilde{\rho}_i\right), \tag{100}$$

where $1 \geq \tilde{\rho}_1 \geq \ldots \geq \tilde{\rho}_s \geq 0$ solves the eigenvalue problem

$$\left|\rho H' S_{11.A_\perp} H - H' S_{10.A_\perp} A\left(A' S_{00.A_\perp} A\right)^{-1} A' S_{01.A_\perp} H\right| = 0, \tag{101}$$

yields

$$\left|A'S_{00.A_\perp.H}A\right| = \left|A'S_{00.A_\perp}A\right|\prod_{i=1}^{s}\left(1-\tilde{\rho}_i\right) \tag{102}$$

This gives the maximized likelihood function, apart from a constant,

$$L_{\max}^{-2/T} = \frac{\left|A_\perp'S_{00}A_\perp\right|\left|A'S_{00.A_\perp}A\right|}{\left|A_\perp'A_\perp\right|\left|A'A\right|}\prod_{i=1}^{r-s}\left(1-\tilde{\lambda}_i\right)\prod_{i=1}^{s}\left(1-\tilde{\rho}_i\right). \tag{103}$$

If C is a $p\times p$ matrix of full rank and $X=\left[A, A_\perp\right]$, where A and $A_\perp$ are full column rank $p\times r$ and $p\times(p\text{-}r)$ matrices respectively, one may use the properties of determinants to write $\left|C\right| = \left|X'CX\right|/\left|X'X\right|$. And since

$$\left|X'X\right| = \begin{vmatrix} A'A & A'A_\perp \\ A_\perp'A & A_\perp'A_\perp \end{vmatrix} = \begin{vmatrix} A'A & 0 \\ 0 & A_\perp'A_\perp \end{vmatrix} = \left|A'A\right|\left|A_\perp'A_\perp\right| \tag{104}$$

$$\left|X'CX\right| = \begin{vmatrix} A'CA & A'CA_\perp \\ A_\perp'CA & A_\perp'CA_\perp \end{vmatrix} = \left|A_\perp'CA_\perp\right|\left|A'CA - A'CA_\perp\left(A_\perp'CA_\perp\right)^{-1}A_\perp'CA\right| \tag{105}$$

we have

$$\left|C\right| = \frac{\left|A_\perp'CA_\perp\right|\left|A'CA - A'CA_\perp\left(A_\perp'CA_\perp\right)^{-1}A_\perp'CA\right|}{\left|A_\perp'A_\perp\right|\left|A'A\right|}. \tag{106}$$

Substituting $S_{00}=C$ and recalling $S_{00.A_\perp} = S_{00} - S_{00}A_\perp\left(A_\perp'S_{00}A_\perp\right)^{-1}A_\perp'S_{00}$

we have

$$\left|S_{00}\right| = \frac{\left|A_\perp'S_{00}A_\perp\right|\left|A'S_{00.A_\perp}A\right|}{\left|A_\perp'A_\perp\right|\left|A'A\right|}. \tag{107}$$

Therefore, apart from a constant, the maximized likelihood is

$$L_{\max}^{-2/T} = \left|S_{00}\right|\prod_{i=1}^{r-s}\left(1-\tilde{\lambda}_i\right)\prod_{i=1}^{s}\left(1-\tilde{\rho}_i\right). \square \tag{108}$$

*Proof of Theorem 2***.** The likelihood ratio test for H_0 in $H(r)$ is

$$LR\left(H_0\left|H(r)\right.\right) = -2\ln\left(L\left(H(r)\right)/L\left(H_0\right)\right). \tag{109}$$

The constant terms in both cancel, and one can write from (15) and (108),

$$LR\left(H_0\left|H(r)\right.\right) = T\left\{\ln\left|S_{00}\right| + \sum_{i=1}^{r-s}\ln\left(1-\tilde{\lambda}_i\right) + \sum_{i=1}^{s}\ln\left(1-\tilde{\rho}_i\right) - \ln\left|S_{00}\right| - \sum_{i=1}^{r}\ln\left(1-\hat{\lambda}_i\right)\right\}, \tag{110}$$

which yields the likelihood ratio test statistic

$$LR\left(H_0\left|H(r)\right.\right) = T\left\{\sum_{i=1}^{r-s}\ln\left(1-\tilde{\lambda}_i\right) + \sum_{i=1}^{s}\ln\left(1-\tilde{\rho}_i\right) - \sum_{i=1}^{r}\ln\left(1-\hat{\lambda}_i\right)\right\}. \tag{111}$$

The calculation for the limiting distribution and degrees of freedom for these tests are based on [20, 21] and on [23, Lemma 7.1]. The former set shows that the limiting distribution of the likelihood ratio tests for restrictions on β and α given r cointegrating relationships is χ^2. The latter shows that for $a\times b$ and $c\times b$ matrices of full column rank, X and Y, the tangent space of XY' has dimension $(a+c\text{-}b)b$. The number of parameters in the unrestricted $\Pi=\alpha\beta'$ is (using, $p=a$, $b=r$, and $c=p$), $2pr\text{-}r^2$. In the restricted model the number of free parameters in $\Pi = A\psi_1H' + A\psi_2\phi'\bar{H}_\perp'$ is $ms+(m+(p\text{-}s)\text{-}(r\text{-}s))(r\text{-}s) = mr+pr\text{-}ps+sr\text{-}r^2$. The

difference between the unrestricted and restricted free parameters, $r(p-m)+s(p-r)$, are the degrees of freedom. So, the likelihood ratio test is asymptotically distributed as χ^2 with $r(p-m)+s(p-r)$ degrees of freedom. □

Proof of Theorem 3. $H_0 : \beta = H\phi,\ \alpha = \left[A, \bar{A}_\perp \psi \right]$ where H $p\times s$, A $p\times m$ are known and ϕ $s\times r$, ψ $(p-m)\times(r-m)$ *are unknown,* $r \leq s < p$.
The reduced rank regression is then

$$R_{0t} = \bar{A}\phi_1' H' R_{1t} + A_\perp \psi \phi_2' H' R_{1t} + \hat{\varepsilon}_t, \tag{112}$$

where φ is partitioned conformably with α as $1 \geq \tilde{\rho}_1 \geq \ldots \geq \tilde{\rho}_s \geq 0$, and is split into

$$\bar{A}' R_{0t} = \phi_1' H' R_{1t} + \bar{A}' \hat{\varepsilon}_t \tag{113}$$

and

$$A_\perp' R_{0t} = \psi \phi_2' H' R_{1t} + A_\perp' \hat{\varepsilon}_t . \tag{114}$$

This allows one to factor the likelihood function into a marginal part based on (114) and a factor based on the (113) conditional on (114):

$$\bar{A}' R_{0t} = \phi_1' H' R_{1t} + \omega \left(A_\perp' R_{0t} - \psi \phi_2' H' R_{1t} \right) + \bar{A}' \hat{\varepsilon}_t - \omega A_\perp' \hat{\varepsilon}_t \tag{115}$$

where

$$\omega = \Omega_{AA_\perp} \Omega_{A_\perp A_\perp}^{-1} = \bar{A}' \Omega A_\perp \left(A_\perp' \Omega A_\perp \right)^{-1}.$$

The parameters in (115) are variation independent of (114) with independent errors.

To maximize the likelihood for the marginal distribution, first one fixes $\hat{\omega}(\varphi_1, \varphi_2, \psi) = \left(\bar{A}' S_{0k} - \varphi_1' H' S_{1k} \right) S_{kk}^{-1}$ in (114) and estimates ψ by regression, giving

$$\hat{\psi} = A_\perp' S_{01} H \phi_2 \left(\phi_2' H' S_{11} H \phi_2 \right)^{-1} \tag{116}$$

and

$$A_\perp' \hat{\varepsilon}_t = A_\perp' R_{0t} - A_\perp' S_{01} H \phi_2 \left(\phi_2' H' S_{11} H \phi_2 \right)^{-1} \phi_2' H' R_{1t} \tag{117}$$

which gives the maximum likelihood estimator for $\Omega_{A_\perp A_\perp} = A_\perp' \Omega A_\perp$,

$$\hat{\Omega}_{A_\perp A_\perp} = A_\perp' S_{00} A_\perp - A_\perp' S_{01} H \phi_2 \left(\phi_2' H' S_{11} H \phi_2 \right)^{-1} \phi_2' H' S_{10} A_\perp . \tag{118}$$

The contribution of the marginal distribution, apart from a constant, is

$$L_{\max M}^{-2/T} = \frac{\left| A_\perp' S_{00} A_\perp - A_\perp' S_{01} H \phi_2 \left(\phi_2' H' S_{11} H \phi_2 \right)^{-1} \phi_2' H' S_{10} A_\perp \right|}{\left| A_\perp' A_\perp \right|} \tag{119}$$

$$= \frac{\left| A_\perp' S_{00} A_\perp \right|}{\left| A_\perp' A_\perp \right|} \frac{\left| \phi_2' H' S_{11} H \phi_2 - \phi_2' H' S_{10} A_\perp \left(A_\perp' S_{00} A_\perp \right)^{-1} A_\perp' S_{01} H \phi_2 \right|}{\left| \phi_2' H' S_{11} H \phi_2 \right|}. \tag{120}$$

One maximizes the factor of the marginal contribution by minimizing the second factor in (120) with respect to the unknown parameter matrix $\hat{\omega}(\varphi_1, \varphi_2, \psi) = \left(\bar{A}' S_{0k} - \varphi_1' H' S_{1k} \right) S_{kk}^{-1}$. This is done by solving the eigenvalue problem

$$\left| \lambda H' S_{11} H - H' S_{10} A_\perp \left(A_\perp' S_{00} A_\perp \right)^{-1} A_\perp' S_{01} H \right| = 0 \tag{121}$$

for eigenvalues $1 \geq \tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_s \geq 0$ and corresponding eigenvectors $\tilde{V} = (\tilde{v}_1, \ldots, \tilde{v}_s)$, normalized so that $\tilde{V}'H'S_{11}H\tilde{V} = I_s$. This implies the maximand of the marginal distribution of the likelihood function is $\hat{\phi}_2 = (\tilde{v}_1, \ldots, \tilde{v}_{r-m})$, from which one then can recover the parameters (40) to (42), and the maximized contribution is

$$L_{\max M}^{-2/T} = \frac{|A_\perp' S_{00} A_\perp|}{|A_\perp' A_\perp|} \prod_{i=1}^{r-m} (1 - \tilde{\lambda}_i). \tag{122}$$

To calculate the conditional distribution, given ϕ_1, $\hat{\phi}_2$, and $\hat{\psi}$, one regresses $\bar{A}'R_{0t} - \phi_1'H'R_{1t}$ on $R_{kt} = A_\perp' R_{0t} - \hat{\psi}\hat{\phi}_2' H' R_{1t}$ to estimate

$$\hat{\omega}(\phi_1, \hat{\phi}_2, \hat{\psi}) = (A_\perp' S_{0k} - \phi_1' H' S_{1k}) S_{kk}^{-1}, \tag{123}$$

where $S_{1k} = \frac{1}{T}\sum_{t=1}^{T} R_{1t}R_{kt}'$ and so on. This allows one to correct for ω in (115) by forming new residual vectors

$$R_{it.k} = R_{it} - S_{ik}S_{kk}^{-1}R_{kt}, \; i = 0,1,k \tag{124}$$

and product moment matrices

$$\begin{aligned} S_{ij.k} &= \frac{1}{T}\sum_{t=1}^{T} R_{it.k}R_{jt.k}' \\ &= S_{ij} - S_{ik}S_{kk}^{-1}S_{kj}, \; i, j = 0,1,k \end{aligned}. \tag{125}$$

One can then write (115) as

$$\bar{A}'R_{0t.k} = \phi_1' H' R_{1t.k} + \hat{u}_t, \tag{126}$$

where $\hat{u}_t = \bar{A}'\hat{\varepsilon}_t - \hat{\omega}A_\perp'\hat{\varepsilon}_t$ and, as suggested by Johansen [20], estimate ϕ_1 by regression which yields

$$\hat{\phi}_1 = (H'S_{11.k}H)^{-1} H'S_{10.k}\bar{A}. \tag{127}$$

The factor of the maximized likelihood corresponding to the conditional distribution is, apart from a constant,

$$L_{\max C}^{-2/T} = \frac{|\hat{\Omega}_{AA.A_\perp}|}{|\bar{A}'\bar{A}|} \tag{128}$$

where

$$\begin{aligned} \Omega_{AA.A_\perp} &= \Omega_{AA} - \Omega_{AA_\perp}\Omega_{A_\perp A_\perp}^{-1}\Omega_{A_\perp A} \\ &= \bar{A}'\Omega\bar{A}' - \bar{A}'\Omega A_\perp (A_\perp'\Omega A_\perp)^{-1} A_\perp'\Omega\bar{A}' \end{aligned}. \tag{129}$$

The maximum likelihood estimate of the conditional variance matrix is

$$\begin{aligned} \hat{\Omega}_{AA.A_\perp} &= \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t' \\ &= \bar{A}'S_{00.k}\bar{A} - \bar{A}'S_{01.k}H(H'S_{11.k}H)^{-1}H'S_{10.k}\bar{A} \end{aligned} \tag{130}$$

which gives, apart from a constant, the maximized likelihood for the conditional distribution as

$$L_{\max C}^{-2/T} = \frac{\left|A'S_{00.k}A - A'S_{01.k}H\left(H'S_{11.k}H\right)^{-1}H'S_{10.k}A\right|}{\left|A'A\right|} \tag{131}$$

$$= \frac{\left|A'S_{00.k}A\right|\left|H'S_{11.k}H - H'S_{10.k}A\left(A'S_{00.k}A\right)^{-1}A'S_{01.k}H\right|}{\left|A'A\right|\left|H'S_{11.k}H\right|} \tag{132}$$

$$= \frac{\left|A'S_{00.k}A\right|}{\left|A'A\right|}\prod_{i=1}^{m}\left(1-\tilde{\rho}_i\right), \tag{133}$$

where $1 \geq \tilde{\rho}_1 \geq \ldots \geq \tilde{\rho}_m > \tilde{\rho}_{m+1} = \ldots = \tilde{\rho}_s = 0$ solve the eigenvalue problem

$$\left|\rho H'S_{11.k}H - H'S_{10.k}A\left(A'S_{00.k}A\right)^{-1}A'S_{01.k}H\right| = 0. \tag{134}$$

The variance-covariance matrix is then estimated by

$$\hat{\Omega} = \begin{bmatrix} A & \bar{A}_\perp \end{bmatrix}\begin{bmatrix} \hat{\Omega}_{AA} & \hat{\Omega}_{AA_\perp} \\ \hat{\Omega}_{A_\perp A} & \hat{\Omega}_{A_\perp A_\perp} \end{bmatrix}\begin{bmatrix} A & \bar{A}_\perp \end{bmatrix}', \tag{135}$$

where the estimators of $\Omega_{A_\perp A_\perp}$, $\omega = \Omega_{AA_\perp}\Omega_{A_\perp A_\perp}^{-1}$, and $\Omega_{AA.A_\perp} = \Omega_{AA} - \Omega_{AA_\perp}\Omega_{A_\perp A_\perp}^{-1}\Omega_{A_\perp A}$ are used to recover $\hat{\Omega}_{A_\perp A_\perp}$, $\hat{\Omega}_{AA_\perp} = \hat{\omega}\hat{\Omega}_{A_\perp A_\perp}$, $\hat{\Omega}_{A_\perp A} = \hat{\Omega}'_{AA_\perp}$, and $\hat{\Omega}_{AA} = \hat{\Omega}_{AA.A_\perp} + \hat{\omega}\hat{\Omega}_{A_\perp A}$.
Finally, the maximized likelihood is

$$L_{\max}^{-2/T} = \frac{\left|A_\perp'S_{00}A_\perp\right|}{\left|A_\perp'A_\perp\right|}\frac{\left|A'S_{00.k}A\right|}{\left|A'A\right|}\prod_{i=1}^{r-m}\left(1-\tilde{\lambda}_i\right)\prod_{i=1}^{m}\left(1-\tilde{\rho}_i\right). \square \tag{136}$$

Proof of Theorem 4. The likelihood ratio test for H_0 in $H(r)$ is

$$LR\left(H_0\middle|H(r)\right) = -2\ln\left(L\left(H_0\right)/L\left(H(r)\right)\right). \tag{137}$$

The constant terms in both cancel, and one can write the likelihood ratio test statistic from (15) and (136),

$$LR\left(H_0 \mid H(r)\right) = T\left\{\ln\left[\frac{\left|A'S_{00.k}A\right|\left|A_\perp'S_{00}A_\perp\right|}{\left|A'A\right|\left|A_\perp'A_\perp\right|}\right] - \ln\left|S_{00}\right| \ldots + \sum_{i=1}^{r-m}\ln\left(1-\tilde{\lambda}_i\right) + \sum_{i=1}^{m}\ln\left(1-\tilde{\rho}_i\right) - \sum_{j=1}^{r}\ln\left(1-\hat{\lambda}_j\right)\right\}. \tag{138}$$

The number of free parameters in the unrestricted model for r cointegrating relationships, from Theorem 2, is $2pr\text{-}r^2$. In the restricted model, $\Pi = A\phi_1'H' + \bar{A}_\perp\psi\phi_2'H'$ (see [23, Lemma 7.1]) has $ms+(p\text{-}m+s\text{-}(r\text{-}m))(r\text{-}m)$ free parameters. The degrees of freedom for the likelihood ratio tests is the difference in free parameters between the unrestricted and restricted models, $m(p\text{-}r)+r(p\text{-}s)$. So, the likelihood ratio test is asymptotically distributed as χ^2 with $m(p\text{-}r)+r(p\text{-}s)$ degrees of freedom. $\square$

Proof of Theorem 5. Under the hypothesis H_0 : $\beta = \left[H, \bar{H}_\perp \phi \right]$, $\alpha = \left[A, \bar{A}_\perp \psi \right]$ where H, A are known $p \times s$ matrices and ϕ and ψ $(p\text{-}s)\times(r\text{-}s)$ are unknown, $s \le r < p$.
The reduced rank regression given H_0 can be expressed as

$$R_{0t} = AH'R_{1t} + \bar{A}_\perp \psi \phi' \bar{H}'_\perp R_{1t} + \varepsilon_t \tag{139}$$

After defining $R_{kt} = R_{0t} - AH'R_{1t}$, for which there are no unknown parameters, one rewrites (139) as

$$R_{kt} = \bar{A}_\perp \psi \phi' \bar{H}'_\perp R_{1t} + \varepsilon_t \tag{140}$$

and premultiplies (140), in turn, by A' and $A'_\perp$ to get

$$A'R_{kt} = A'\varepsilon_t \tag{141}$$

and

$$A'_\perp R_{kt} = \psi_1 \phi' \bar{H}'_\perp R_{1t} + A'\varepsilon_t . \tag{142}$$

This allows one to factor the likelihood function into a marginal part based on (141) and a factor based on (142) conditional on (141):

$$A'_\perp R_{kt} = \psi \phi' \bar{H}'_\perp R_{1t} + \omega A' R_{kt} + A'_\perp \hat{\varepsilon}_t - \omega A' \hat{\varepsilon}_t , \tag{143}$$

where $\omega = \Omega_{A_\perp A} \Omega_{AA}^{-1} = \bar{A}'_\perp \Omega A \left(A' \Omega A \right)^{-1}$. The parameters in (143) are variation independent of (141) with independent errors.
To calculate the conditional factor, one fixes ϕ and ψ and regresses $A'_\perp R_{kt} - \psi \phi' \bar{H}'_\perp R_{1t}$ on $\omega A' R_{kt}$ to estimate

$$\omega(\phi, \psi) = \left(A'_\perp S_{kk} A - \psi \phi' \bar{H}'_\perp S_{1k} A \right) \left(A' S_{kk} A \right)^{-1} \tag{144}$$

This allows one to correct for ω in (143) by forming new residual vectors

$$R_{it.A} = R_{it} - S_{ik} A \left(A' S_{kk} A \right)^{-1} A' R_{kt}, \; i = 1, k \tag{145}$$

and product moment matrices

$$\begin{aligned} S_{ij.A} &= \frac{1}{T} \sum_{t=1}^{T} R_{it.A} R'_{jt.A} \\ &= S_{ij} - S_{ik} A \left(A' S_{kk} A \right)^{-1} A' S_{kj}, \; i, j = 1, k \end{aligned} . \tag{146}$$

This allows one to write (143) as

$$A'_\perp R_{kt.A} = \psi \phi' \bar{H}'_\perp R_{1t.A} + \hat{u}_t , \tag{147}$$

where $\hat{u}_t = A'_\perp \hat{\varepsilon}_t - \hat{\omega} A' \hat{\varepsilon}_t$. Fixing ϕ, one estimates ψ by regressing $A'_\perp R_{kt.A}$ on $\phi' \bar{H}'_\perp R_{1t.A}$, which yields

$$\hat{\psi}(\phi) = A'_\perp S_{k1.A} \bar{H}_\perp \phi \left(\phi' \bar{H}'_\perp S_{11.A} \bar{H}_\perp \phi \right)^{-1} . \tag{148}$$

The factor of the maximized likelihood corresponding to the conditional distribution is, apart from a constant,

$$L_{\max C}^{-2/T} = \frac{\left| \hat{\Omega}_{A_\perp A_\perp . A} \right|}{\left| A'_\perp A_\perp \right|}, \tag{149}$$

where

$$\begin{aligned}\Omega_{A_\perp A_\perp .A} &= \Omega_{A_\perp A_\perp} - \Omega_{A_\perp A}\Omega_{AA}^{-1}\Omega_{AA_\perp} \\ &= A_\perp'\Omega A_\perp - A_\perp'\Omega A\left(A'\Omega A\right)^{-1}A'\Omega A_\perp\end{aligned}. \tag{150}$$

The maximum likelihood estimate of the conditional variance matrix is

$$\begin{aligned}\hat{\Omega}_{A_\perp A_\perp .A} &= \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t\hat{u}_t' \\ &= A_\perp' S_{kk.A}A_\perp - A_\perp' S_{k1.A}\bar{H}_\perp\phi\left(\phi'\bar{H}_\perp' S_{11.A}\bar{H}_\perp\phi\right)^{-1}\phi'\bar{H}_\perp' S_{1k.A}A_\perp\end{aligned},$$

which gives, apart from a constant, the maximized likelihood for the conditional factor as

$$L_{\max C}^{-2/T} = \frac{\left|A_\perp' S_{kk.A}A_\perp - A_\perp' S_{k1.A}\bar{H}_\perp\phi\left(\phi'\bar{H}_\perp' S_{11.A}\bar{H}_\perp\phi\right)^{-1}\phi'\bar{H}_\perp' S_{1k.A}A_\perp\right|}{\left|A_\perp' A_\perp\right|} \tag{151}$$

$$= \frac{\left|A_\perp' S_{kk.A}A_\perp\right|\left|\phi'\bar{H}_\perp' S_{11.A}\bar{H}_\perp\phi - \phi'\bar{H}_\perp' S_{1k.A}A_\perp\left(A_\perp' S_{kk.A}A_\perp\right)^{-1}A_\perp' S_{k1.A}\bar{H}_\perp\phi\right|}{\left|A_\perp' A_\perp\right|\left|\phi'\bar{H}_\perp' S_{11.A}\bar{H}_\perp\phi\right|}. \tag{152}$$

The conditional likelihood is maximized by minimizing (152) with respect to ϕ, which is done by solving the eigenvalue problem

$$\left|\lambda\bar{H}_\perp' S_{11.A}\bar{H}_\perp - \bar{H}_\perp' S_{1k.A}A_\perp\left(A_\perp S_{kk.A}A_\perp\right)^{-1}A_\perp' S_{k1.A}\bar{H}_\perp\right| = 0 \tag{153}$$

for $1 \geq \tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_{s-r} \geq \tilde{\lambda}_{s-r+1} = \ldots = \tilde{\lambda}_{p-s} = 0$ and for eigenvectors $\tilde{V} = \left(\tilde{v}_1, \ldots, \tilde{v}_{p-s}\right)$, normalized so that $\tilde{V}'\bar{H}_\perp' S_{11.A}\bar{H}_\perp\tilde{V} = I_{p-r}$. The maximand of the likelihood function is $\hat{\phi} = \left(\tilde{v}_1, \ldots, \tilde{v}_{r-s}\right)$, from which one then can recover the parameters (52) to (54), and the maximized likelihood function for the conditional piece, apart from a constant, is

$$L_{\max C}^{-2/T} = \frac{\left|A_\perp' S_{kk.A}A_\perp\right|}{\left|A_\perp' A_\perp\right|}\prod_{i=1}^{r-s}\left(1-\tilde{\lambda}_i\right). \tag{154}$$

The maximum of the factor corresponding to the likelihood function for the marginal piece based on (141) is, apart from a constant, $L_{\max M}^{-2/T} = \frac{\left|\hat{\Omega}_{AA}\right|}{\left|A'A\right|}$. The denominator is estimated by

$\hat{\Omega}_{AA} = A'\hat{\Omega}A = \frac{1}{T}\left(A'\hat{\Omega}A\right) = \frac{1}{T}\left(A'\hat{\varepsilon}\hat{\varepsilon}'A\right) = \frac{1}{T}\left(A'R_kR_k'A\right) = A'S_{kk}A$, and thus

$$L_{\max M}^{-2/T} = \frac{\left|A'S_{kk}A\right|}{\left|A'A\right|}. \tag{155}$$

The variance-covariance matrix is then estimated by

$$\hat{\Omega} = \begin{bmatrix}A & A_\perp\end{bmatrix}\begin{bmatrix}\hat{\Omega}_{AA} & \hat{\Omega}_{AA_\perp} \\ \hat{\Omega}_{A_\perp A} & \hat{\Omega}_{A_\perp A_\perp}\end{bmatrix}\begin{bmatrix}A & A_\perp\end{bmatrix}', \tag{156}$$

where the estimators of Ω_{AA}, $\omega = \Omega_{A_\perp A}\Omega_{AA}^{-1}$, and $\Omega_{A_\perp A_\perp .A} = \Omega_{A_\perp A_\perp} - \Omega_{A_\perp A}\Omega_{AA}^{-1}\Omega_{AA_\perp}$ are used to recover $\hat{\Omega}_{AA}$, $\hat{\Omega}_{A_\perp A} = \hat{\omega}\hat{\Omega}_{AA}$, $\hat{\Omega}_{AA_\perp} = \hat{\Omega}_{A_\perp A}'$, and $\hat{\Omega}_{A_\perp A_\perp} = \hat{\Omega}_{A_\perp A_\perp .A} + \hat{\omega}\hat{\Omega}_{AA_\perp}$.

The product of (154) and (155) yield, apart from a constant, the maximized likelihood

$$L_{\max}^{-2/T} = \frac{\left|A'S_{kk}A\right|\left|A_\perp' S_{kk.A} A_\perp\right|}{\left|A'A\right|\left|A_\perp' A_\perp\right|} \prod_{i=1}^{r-s}\left(1-\tilde{\lambda}_i\right). \tag{157}$$

By the same arguments used in (104) to (107), one can show that

$$\left|S_{kk}\right| = \frac{\left|A'S_{kk}A\right|\left|A_\perp' S_{kk.A} A_\perp\right|}{\left|A'A\right|\left|A_\perp' A_\perp\right|} \tag{158}$$

so that the maximized likelihood function is

$$L_{\max}^{-2/T} = \left|S_{kk}\right| \prod_{i=1}^{r-s}\left(1-\tilde{\lambda}_i\right). \square \tag{159}$$

Proof Theorem 6. The likelihood ratio test for H_0 in $H(r)$ is

$$LR\left(H_0 \middle| H(r)\right) = -2\ln\left(L(H_0)/L(H(r))\right). \tag{160}$$

The constant terms in both cancel, and one can write the likelihood ratio test statistic from (15) and (159),

$$LR\left(H_0 \mid H(r)\right) = T\left\{\ln\left|S_{kk}\right| - \ln\left|S_{00}\right| + \sum_{i=1}^{r-s}\ln\left(1-\tilde{\lambda}_i\right) - \sum_{j=1}^{r}\ln\left(1-\hat{\lambda}_j\right)\right\}. \tag{161}$$

The number of free parameters in the unrestricted model for r cointegrating relationships, from Theorem 2, is $2pr - r^2$. In the restricted model, $\Pi = AH' + \bar{A}_\perp \psi \phi' H'$ (from [23, Lemma 7.1]) has $((p-s)+(p-s)-(r-s))(r-s)$ free parameters. The degrees of freedom for the likelihood ratio tests is the difference in free parameters between the unrestricted and restricted models, $s(2p-s)$. So, the likelihood ratio test is asymptotically distributed as χ^2 with $2ps - s^2$ degrees of freedom. $\square$

Proof of Theorem 7. $H_0 : \beta = H\phi$ where H $p\times s$ is known and ϕ $s\times r$ is unknown, $r \le s < p$. That one may choose $\beta_\perp = \left[H_\perp, \bar{H}\phi_\perp\right]$ as the orthogonal complement of β was shown is section 4.
Consider

$$\beta_\perp = \left[G, \bar{G}_\perp \theta\right] \tag{162}$$

where G $p\times q$ is known and θ $(p-q)\times(p-r)$ is unknown. $\beta_\perp = \left[G, \bar{G}_\perp \theta\right]$ implies $sp(G) \subset sp(\beta_\perp)$, which implies $sp(\beta) \subset sp(G_\perp)$. Setting $H = G_\perp$, a $p\times(p-q)$ matrix, implies $sp(\beta) \subset sp(H)$, which shows this is a test of the form $\beta = H\phi$ where ϕ is $(p-q)\times r$. Noting $\beta_\perp'\beta = \left[G, \bar{G}_\perp\theta\right]' H\phi = \begin{bmatrix} G'H\phi \\ \theta'\bar{G}_\perp' H\phi \end{bmatrix} = \begin{bmatrix} 0\phi \\ \theta' I_{p-q}\phi \end{bmatrix}$ is zero when $\theta = \phi_\perp$ (that is, $\phi = \theta_\perp$) and setting $s = q-p$ shows this is test (1) in section 3. $\square$

Proof of Theorem 8. $H_0 : \beta = \left[H, \bar{H}_\perp \phi\right]$ where H $p\times s$ is known and ϕ $(p-s)\times(s-r)$ is unknown. That one may choose $\beta_\perp = H_\perp \phi_\perp$ was shown in section 4. Consider

$$\beta_\perp = G\theta \tag{163}$$

where G is a known $p \times q$ matrix and θ is an unknown $q \times (p-r)$ matrix; $\beta_\perp = G\theta$ implies $sp(\beta_\perp) \subset sp(G)$, which implies $sp(G_\perp) \subset sp(\beta)$. Setting the $H = G_\perp$, a $p \times (p-q)$ matrix, implies $sp(H) \subset sp(\beta)$, which shows this is a test of the form $\beta = [H, \zeta]$ where ζ is $r \times (p-q)$. Noting $\beta_\perp'\beta = \theta' G' [H, \zeta] = [\theta' G_\perp' H, \theta' G' \zeta] = [\phi_\perp' 0, \theta' G' \zeta] = \left[0_{(p-r)\times p-q}, \theta' G' \zeta\right]$ is zero when $\zeta = G_\perp \left(G_\perp' G_\perp\right)^{-1} \phi_\perp$ and setting $s=q-p$ shows this is test (1) in section 3. □

The other theorems in section 4 are combinations of the above theorems using α and β.