Structural Breaks in Cointegrated VAR Models

Håvard Hungnes*

June 30, 2003

Abstract

We describe a procedure for decomposing the deterministic terms into growth rate parameters and cointegration mean parameters. These parameters express long-run properties of the model. For example, the growth rate parameters tell us how much to expect (unconditionally) the variables in the system to grow from one period to the next, and will therefore be especially important to identify if the model is to be used for forecasting.

The procedure can be used for analysing structural breaks when the deterministic terms include shift dummies. By decomposing the coefficients into interpretable components, different types of structural breaks can be identified. Both shifts in intercepts and shifts in growth rates, or combinations of these, can be tested for. The ability to distinguish between different types of structural shifts makes the procedure superior compared to alternative procedures. Furthermore, the procedure utilizes the information more efficiently than alternative procedures, where observations are disregarded (Johansen et al., 2000, EcJ) or two-step de-trending is used (Saikkonen & Lütkepohl, 2000, JBES).

The procedure is programmed in Ox.

Work in Progress!

Keywords: Johansen procedure, cointegrated VAR, structural shifts, growth rates, cointegration means.

JEL classification: C32, C51, C52.

Acknowledgments: Thanks to Søren Johansen, Ragnar Nymoen and Anders Rygh Swensen for valuable comments on earlier versions of this paper.

*E-mail: havard.hungnes@ssb.no Homepage: http://folk.ssb.no/hhu

1 Introduction

In analysing a dynamic economic model we are often interested in identifying and testing its long run properties. The cointegrating vectors are examples of long run relationships between different data series. However, also the underlying growth rates (i.e. steady state growth rates) can be identified in cointegrated vector autoregressive (VAR) models. Hungnes (2002) shows how these growth rates can be estimated in a full information maximum likelihood framework, as well as how to test for restrictions on these growth rates.

The growth rates tell us how much to expect (unconditionally) the variables in the system to grow from one period to the next. If the system is used for forecasting, the vector of growth rates will be one of the most important ones in providing good forecasts. In fact, as the forecasting horizon approaches infinity, the forecast will rely on this vector only.

Structural breaks often imply changes in the growth rates of the variables. With many and frequent structural breaks in time series integrated of order 1, it will normally be best to estimate the system as if it were integrated of order 2. With less frequent structural changes these changes can be identified.

Structural breaks have been discussed intensively in the context of univariate autoregressive time series. Perron (1989) suggests three models: Model A, a 'crash model', with change in intercept but where the slope of the linear trend is unchanged; Model B, a 'changing growth model', allows a change in the slope of trend function without any sudden change in the level at the time of the break; And model C, where both intercept and slope are changed at the time of the break. Johansen et al. (2000) present a generalization of model C in a multivariate framework, and allow for testing hypothesis corresponding to model A.

In this paper we decompose all deterministic terms in a cointegrated VAR model into interpretable counterparts. The corresponding coefficients describe the long run (steady state) growth rates in the variables, and possibly shifts in level and growth rates (the latter depending on the type of deterministic variables that are included in the system). Combined with the coefficients for the cointegrating vectors, they also describe level and trends (and possibly shifts in these) in the cointegrating vectors. This decomposition therefore allows us to test all three types of structural breaks suggested in Perron (1989). As Johansen et al. (2000) we present a model C for a multivariate framework, but where we allow for testing hypothesis corresponding to both A and B. I addition the method presented here makes it possible to identify the growth rates and the size of the different types of shifts.

Another advantage of the method presented here compared to the one in Johansen et al. (2000), is that we make use of all the information. Johansen et al. (2000) disregards the observations following right after the break by including impulse dummies. In Johansen et al (2000) the number of impulse dummies after the break corresponds to the number of lags in the system, which implies a reduction in the degrees of freedom compared with our approach.

Saikkonen and Lütkepohl (2000) suggest a two step approach to estimate cointegrated VAR models with structural breaks. In the first step all the coefficients for the deterministic variables are estimated. In the second step a normal cointegration analysis is conducted, but where the deterministic components are removed from each time series. The estimation in the second step is therefore done without any deterministic variables included. One problem with this estimation method is that not all restrictions among the coefficients for the deterministic variables can be taken into account in the first step. In the same way as for the estimation procedure in Johansen et al. (2000) this involves a reduction in the degrees of freedom when the coefficients for the deterministic variables are estimated.

We will illustrate our new estimation procedure presented on two data sets with structural breaks. However, we think it is important to stress that this is only one type of problem this method can be used to analyse. A decomposition of the deterministic terms into interpretable counterparts will often be important even though there are no structural breaks. As mentioned above, the growth rates are important in forecasting, and identifying the growth rates is important to judge the forecasting ability of a model.

Another application for the procedure is to test for "zero-mean" convergence. "Zero-mean" convergence implies that the difference between two variables (say; output per capita in two countries) is stationary with zero mean. To test if the mean is zero, the deterministic terms must be decomposed.

In a simpler version of the estimation procedure, Hungnes (2002) used it to estimate a cointegrated VAR model were some variables where allowed to grow and some were not.

Throughout the paper we define the orthogonal complement of the full column rank matrix A as $A_{\perp}$ such that $A'_{\perp}A = 0$ and $(A, A_{\perp})$ has full rank. (The orthogonal complement of a nonsingular matrix is 0, and the orthogonal complement of a zero matrix is an identity matrix of suitable dimensions.)[1] Further, let A^{+} be the Moore-Penrose inverse of A, see e.g. Theil (1983, pp. 51-52) or Magnus and Neudecker (1988, pp. 32-34).[2] If A has full column rank the transpose of the Moore-Penrose inverse is given by $\overline{A} = A\left(A'A\right)^{-1}$.

2 Model formulation without structural breaks

2.1 Normal representation

In (1) Y_t is an n-dimensional vector of variables that are integrated of order one at most. α and β are matrixes of dimension $n \times r$ (where r is the number of cointegration vectors) and $\beta' Y_t$ is $I(0)$. Furthermore, Γ_i is an $n \times n$ matrix of coefficients and Δ is the difference operator. D_t is a vector of deterministic variables. The errors ε_t are

[1] If the rank of the $n_1 \times n_2$ matrix A is $n^* < n_2$, then $A_{\perp}$ is an $n_1 \times (n_1 - n^*)$ matrix such that $A'_{\perp}A = 0$ and $(A_{\perp}, A)$ has full row rank if $n^* < n_1$, and $A_{\perp} = 0$ if $n^* = n_1$.

[2] The Moore-Penrose matrix A^{+} satifies the following requirements; $AA^{+}A = A$, $A^{+}AA^{+} = A^{+}$, and both $A^{+}A$ and AA^{+} are symmetric.

Table 1: Deterministic part - normal representation

	$\delta D_t = \delta_0 + \delta_1 t$		...		...	
	$\delta_0 =$	$\delta_1 =$	Y	$\beta' Y$	constant	trend
H_{ql}	$\alpha\mu + \alpha_\perp \mu^*$	$\alpha\rho + \alpha_\perp \rho^*$	quadratic	linear	unrestricted	unrestricted
H_l	$\alpha\mu + \alpha_\perp \mu^*$	$\alpha\rho$	linear	linear	unrestricted	restricted
H_{lc}	$\alpha\mu + \alpha_\perp \mu^*$	0	linear	constant	unrestricted	absent
H_c	$\alpha\mu$	0	constant	constant	restricted	absent
H_{cz}	0	0	constant	zero	absent	absent

assumed to be white noise Gaussian ($\varepsilon_t \sim N(0, \Omega)$).

$$\Delta Y_t = \alpha (\beta' Y_{t-1}) + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \delta D_t + \varepsilon_t, \quad t = 1, 2, ..., T. \tag{1}$$

It is common to distinguish between deterministic variables that are restricted to lie in the cointegration space and those which are not. Table 1 summarises the normal specifications where a constant and possibly a trend is included. Let $\delta D_t = \delta_0 + \delta_1 t$, where δ_0 and δ_1 are vectors of dimension n. Furthermore, let $\delta_0 = \alpha\mu + \alpha_\perp \mu^*$ and $\delta_1 = \alpha\rho + \alpha_\perp \rho^*$, where μ and ρ are vectors of dimension r and μ^* and ρ^* are vectors of dimension $n - r$. Then μ represents the part of the constant restricted to lie in the cointegration space. Therefore, if $\delta_0 = \alpha\mu$ the constant is restricted to lie in the cointegration space.

In, say, row H_l the constant is included unrestricted whereas the trend is restricted to lie in the cointegration space. This implies that the variables Y will follow linear trends, and we also allow the linear combination of the variables represented by the cointegration vectors $\beta' Y$ to follow linear trends. If the trend were not restricted to lie in the cointegration space, we would allow the variables in our system to follow a quadratic trend. This is not realistic for most macroeconomic data series (especially when we analyse the logarithm of the series), and is therefore hardly ever used in econometric analysis.

We assume that the process in (1) is generated by hypothesis H_l in Table 1. The system grows at the unconditional rate $E_t [\Delta Y_t] = \gamma$ with long run (cointegration) means $E_t [\beta' Y_t] = \mu + \rho t$. We can rewrite the system as

$$\Delta Y_t - \gamma = \alpha (\beta' Y_{t-1} - \mu - \rho (t-1)) + \sum_{i=1}^{p-1} \Gamma_i (\Delta Y_{t-i} - \gamma) + \varepsilon_t, \tag{2}$$

where

$$\beta' \gamma = \rho. \tag{3}$$

Theorem 1 (Grangers representation theorem) *Under the assumption that the system has none of its characteristics roots inside the complex unit circle and $\alpha'_\perp \Gamma \beta_\perp$ has full rank, Y_t has the representation*

$$Y_t = C \sum_{i=1}^{t} \varepsilon_t + \iota + \gamma t + A_t, \tag{4}$$

where $C = \beta_\perp \left(\alpha'_\perp \Gamma \beta_\perp\right)^{-1} \alpha'_\perp$. *The process* A_t *is stationary with zero expectation. The growth rate parameters can be expressed as*

$$\gamma = C\delta_0 - (C\Gamma - I_n)\,\overline{\beta}\rho, \tag{5}$$

whereas the level coefficients ι *depends on initial values in such a way that*

$$\mu = \beta'\iota = \overline{\alpha}\,(\Gamma C - I_n)\,\delta_0 - \overline{\alpha}'\Gamma\,(C\Gamma - I_n)\,\overline{\beta}\rho - \rho. \tag{6}$$

Proof. For proof see Theorem 4.2. in Johansen (1995) or Theorem 2.1 in Johansen et al.(2000). ∎

In econometric analysis it is recommended to start with system H_l or H_c. If there are trends in the data H_l is recommended, and in systems without trends H_c is recommended.

The system in (1) is basically a multivariate version of the Dikey-Fuller test. In the Dickey-Fuller test we test whether a variable follows a stochastic trend (the null hypothesis) or a deterministic trend (the alternative hypothesis). When determining the cointegration rank of the system in (1), we want to identify the number of stochastic trends. In the trace statistics the alternative hypothesis is that the system does not have any stochastic trends. If the trend is included and restricted to lie in the cointegration space, the alternative hypothesis implies that the system has up to n deterministic trends.

In a Monte Carlo simulation Doornik et al. (1998), it is shown that if the system is miss-specified by not including a restricted trend, we will often identify too few cointegration vectors. This is because deterministic trends will be replaced by stochastic trends (p. 133). However, to erroneously include the trend restricted has a very low cost (p. 133).

2.2 Alternative representation

Here we will represent the cointegrated system slightly different. The reason for changing the representation is that it will then be easier to represent the system with structural breaks.

Let X be the part of Y where the deterministic components are removed, i.e.

$$X_t = Y_t - \gamma D_t$$

with D as the vector of deterministic variables and γ as the corresponding matrix of coefficients. The system can therefore alternatively be written as

$$\Delta X_t = \alpha\left(\beta' X_{t-1} - \mu\right) + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \varepsilon_t.$$

We will now consider the two cases H_c and H_l more throughout.

Case 2 (H_c) *Here the intercept in the cointegration space is the only deterministic variable. Therefore;* $X_t = Y_t$. *Then the system is*

$$\Delta Y_t = \alpha\left(\beta' Y_{t-1} - \mu\right) + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t. \tag{7}$$

Case 3 **(H_l)** *In this case there is a trend in the variables, Y, and this trend is removed in X. Therefore; $X_t = Y_t - \gamma t$, i.e. $D_t = t$. The system is*

$$\Delta Y_t - \gamma = \alpha \left(\beta' \left(Y_{t-1} - \gamma \left(t-1\right)\right) - \mu\right) + \sum_{i=1}^{p-1} \Gamma_i \left(\Delta Y_{t-i} - \gamma\right) + \varepsilon_t, \tag{8}$$

where $\rho \equiv \beta'\gamma$ represents the deterministic trends in the cointegration space.

Also other deterministic variables can be included in the system. Centered seasonal dummies are examples of such variables. In the case with linear trends in the system variables (i.e. H_l), $X_t = Y_t - \gamma D_t$ where $D_t = (t, S_{1,t}, S_{2,t}, S_{3,t})$ with quarterly data. Normally centered seasonal dummies will be defined such that $S_{2,t} = S_{1,t-1}$ and $S_{3,t} = S_{2,t-1} = S_{1,t-2}$. For short, we therefore use $D_t = (t, S_t, S_{t-1}, S_{t-2})$.

$$\Delta Y_t - \gamma \Delta D_t = \alpha \left(\beta' \left(Y_{t-1} - \gamma D_{t-1}\right) - \mu\right) + \sum_{i=1}^{p-1} \Gamma_i \left(\Delta Y_{t-i} - \gamma \Delta D_{t-i}\right) + \varepsilon_t \tag{9}$$

where $\beta'\gamma = \rho$ are the trends in the cointegration space..

The variable ΔS_t is a centered seasonal dummy itself since it is only a linear combination of two seasonal dummies; $\Delta S_{1,t} = S_{1,t} - S_{1,t-1} = S_{1,t} - S_{2,t}$. The formulation above therefore corresponds to including seasonal dummies 'unrestricted'. In the formulation above it is possible to identify the seasonal pattern.

Note

$$\begin{pmatrix} \Delta S_{1,t} \\ \Delta S_{2,t} \\ \Delta S_{3,t} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} S_{1,t} \\ S_{2,t} \\ S_{3,t} \end{pmatrix}, \tag{10}$$

because

$$S_{4,t} = -\left(S_{1,t} + S_{2,t} + S_{3,t}\right). \tag{11}$$

By applying this relationship we could transform estimated coefficients for the seasonal dummies into their interpretable counterparts. However, this is complicated, and the method can not be used if we include other deterministic dummies such as step dummies in our system. In the next section we derive a conditional estimator which we can apply in a switching algorithm in order to find the full information maximum likelihood estimates for the coefficients in the system.

3 Model formulation with structural break

If there are structural breaks in the time series, there might be both level shifts and trend shifts. A shift dummy picks up the level shift in a time series. A shift dummy is a dummy equal to zero up till one period and unity afterwords.

A broken trend picks up the trend shift. The broken trend is constructed as the accumulated value of the corresponding shift dummy. The accumulated shift dummy is therefore continuous, see Figure 1.

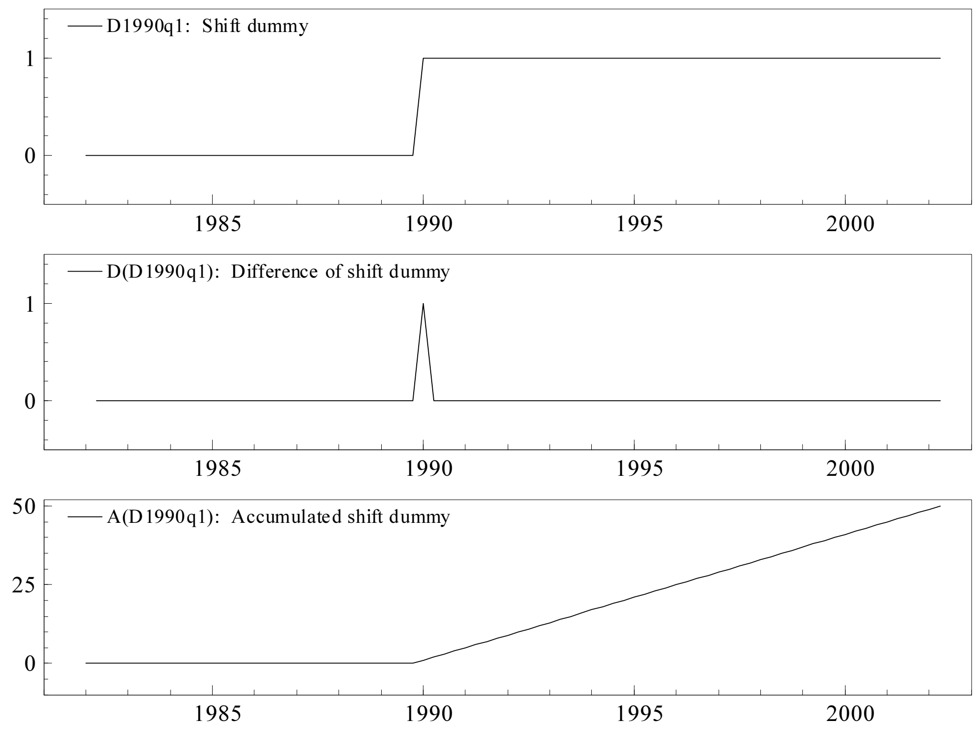

Figure 1: Shift dummy, with corresponding impulse dummy (i.e. the differenciate of the shift dummy) and broken trend (i.e. the accumulated shift dummy).

Both the shift dummy and the corresponding broken trend are included in the vector of deterministic variables, D_t. The general system becomes

$$\Delta Y_t - \gamma \Delta D_t = \alpha\beta' \left((Y_{t-1} - \gamma D_{t-1}) - \mu \right) + \sum_{i=1}^{p-1} \Gamma_i \left[\Delta Y_{t-i} - \gamma \Delta D_{t-i} \right] + \varepsilon_t \qquad (12)$$

where $\beta'\gamma = \rho$. The difference of a shift dummy is an impulse dummy, being equal to unity in the first period the shift dummy is unity and zero otherwise.

In the cointegrating vector a significant coefficient for the shift dummy implies a change in the intercept of the cointegrating vector. A significant coefficient for the accumulated shift dummy (i.e. the broken trend) implies a shift in the slope of the trend in the cointegrating vector. Therefore, according to Perron (1989) the cointegration vector becomes a 'crash model' (A) if the coefficient for the shift dummy is significant whereas the coefficient for the broken trend is not significant. On the other hand, the cointegrating vector behaves as a 'changing growth model' (B) if the coefficient for the shift dummy is not significant whereas the coefficient for the broken trend is significant. If both coefficients are significant the cointegrating vector behaves as C.

If all cointegrating vectors are of type A, then the cointegration space can be said to behave as a 'crash model'. Similarly, if all cointegrating vectors are of type B, we say that the cointegration space behaves as a 'changing growth model'.

A time series follows a 'crash model' (A) if the corresponding coefficient for the difference of the shift dummy (i.e. the shift dummy) is significant whereas the corresponding coefficient for the shift dummy is not significant. And a time series follows a 'changing growth model' (B) if the corresponding coefficient for the difference of the shift dummy is not significant whereas the corresponding coefficient for the shift dummy is significant.

If all time series follows a 'crash model' (A), then it is impossible to construct a linear relationship of the time series of type B. Therefore the coefficients for the broken trend in the cointegrating space must be zero as well. Similarly, if all time series follows a 'changing growth model' (B) none of the cointegrating vectors can be of type A.

However, the implication does not go the other way around. If the cointegrating vectors follow a 'crash model' this does not imply that the time series themselves follow a 'crash model'. The time series may follow a model of type C. If this is the case, we say that the cointegrating vectors correspond to the co-breaking vectors.

3.1 Estimation

3.1.1 Restrictions on γ

By defining $Y_t^* = vec\left(Y_t, 1\right)$ and $\beta^{*\prime} = \left(\beta', -\mu\right)$ the system can be written as

$$\begin{aligned} &\Delta Y_t - \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} \\ = \quad &\alpha\beta^{*\prime} Y_{t-1}^* + \left(I, -\Gamma_1^* - \alpha\beta', -\Gamma_2^*, ..., -\Gamma_p^*\right) \left(I_{p+1} \otimes \gamma\right) \begin{pmatrix} D_t \\ D_{t-1} \\ \vdots \\ D_{t-p} \end{pmatrix} + \varepsilon_t \end{aligned} \tag{13}$$

where $\Gamma_1^* = \Gamma_1 + I_n$, $\Gamma_i^* = \Gamma_i - \Gamma_{i-1}$ $(i = 2, 3, ..., p-1)$ and $\Gamma_p^* = -\Gamma_{p-1}$.

First we investigate the case where restrictions are imposed on γ. The restrictions we want to impose on γ can be written as $R_\gamma' \gamma' = 0$, or equivalently $\left(I_n \otimes R_\gamma'\right) vec\gamma' = 0$. These restrictions can be restrictions on both the change in the intercept in the individual time series, as well as the change in the growth rates.

The total set of restrictions on γ can alternatively be written as

$$vec\gamma' = H_\gamma \psi, \tag{14}$$

where $H_\gamma = \left(I_n \otimes R_\gamma\right)_\perp$.

Define $Z_t = vec\left(\Delta Y_t, ..., \Delta Y_{t-p+1}\right)$, $D_t^* = vec\left(D_t, ..., D_{t-p}\right)$, $\Phi = \left(I_n, -\Gamma_1, ..., \Gamma_{p-1}\right)$, and $\Phi^* = \left(I, -\Gamma_1^* - \alpha\beta', -\Gamma_2^*, -\Gamma_3^*, ..., -\Gamma_p^*\right)$. The cointegrated VAR becomes

$$\Phi Z_t = \Phi^* \left(I_{p+1} \otimes \gamma\right) D_t^* + \alpha\beta^{*\prime} Y_{t-1}^* + \varepsilon_t. \tag{15}$$

Define

$$\begin{aligned} S_{DD} &= T^{-1}\sum D_t^* D_t^{*\prime}, \\ S_{ZD} &= T^{-1}\sum Z_t D_t^{*\prime}, \\ S_{YD} &= T^{-1}\sum Y_{t-1}^* D_t^{*\prime}, \end{aligned}$$

and similarly for S_{ZZ} and $S_{YZ} = S_{ZY}^{\prime}$. Furthermore, we implicitly define the matrix M with dimension $q\,(p+1)^2\, n \times qn$ (where q is the number of deterministic variables in D_t) by[3]

$$vec\,(I_{p+1} \otimes \gamma) = M vec\gamma^{\prime}. \tag{17}$$

We have now the following theorem:

Theorem 4 *With given estimates of* $\alpha, \beta^*, \Gamma_1, ..., \Gamma_{p-1}$ *and* Ω*, the conditional estimator for* γ *under the restriction (14) becomes*

$$\begin{aligned} & \widehat{vec\gamma^{\prime}}\,(\alpha, \beta^*, \Gamma_1, ..., \Gamma_{p-1}, \Omega) \\ = \;& H_\gamma \left[H_\gamma^{\prime} M^{\prime} \left(S_{DD} \otimes \Phi^{*\prime}\Omega^{-1}\Phi^*\right) M H_\gamma^{\prime}\right]^{-1} \\ & \times \left[H_\gamma^{\prime} M^{\prime} vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime} S_{YD}\right)\right]. \end{aligned} \tag{18}$$

Proof. For proof, see the appendix. ■

3.1.2 Restrictions on ρ

As discussed above, restrictions on γ implies restrictions on ρ. However, sometimes we want to impose additional restrictions on ρ. Let these restrictions be written as

$$R_\rho^{\prime}\rho^{\prime} = 0. \tag{19}$$

We will not include ρ directly in our system. Instead we will make use of the relationship between ρ and γ. Therefore we have to transform the restrictions on ρ into restrictions on γ. Since $\beta^{\prime}\gamma = \rho$, the restriction on ρ can be written as $R_\rho^{\prime}\gamma^{\prime}\beta = 0$. By vectorisation, this becomes $\left(\beta^{\prime} \otimes R_\rho^{\prime}\right) vec\gamma^{\prime} = 0$. By combining these restrictions via ρ by those directly imposed on γ, the total set of restrictions on γ therefore becomes

$$\begin{pmatrix} \beta^{\prime} \otimes R_\rho^{\prime} \\ I_n \otimes R_\gamma^{\prime} \end{pmatrix} vec\gamma^{\prime} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which we simplify to $R^{\prime} vec\gamma^{\prime} = 0$. The total set of restrictions on γ can alternatively be written as

$$vec\gamma^{\prime} = H\psi, \tag{20}$$

where $H = R_\perp$.

[3] The explicite form of this matrix is given in e.g. Magnus and Neudecker (1979, pp. 47-48) as

$$M = \left(\left(\left(I_{p+1} \otimes K_{q,p+1}\right)\left(vecI_{p+1} \otimes I_q\right)\right) \otimes I_n\right) K_{q,n}, \tag{16}$$

where $K_{q,n}$ is the commutation matrix, see e.g. Magnus and Neudecker (1979, pp. 46-48). This is defined as $K_{q,n} = \sum_{i=1}^{q}\sum_{j=1}^{n}\left(J_{i,j} \otimes J_{i,j}^{\prime}\right)$, where $J_{i,j}$ is an $q \times n$ matrix with all elements equal to zero except the (i,j)'th element, which is unity (see e.g. Lütkepohl, 1996, p. 116).

Theorem 5 *With given estimates of* $\alpha, \beta^*, \Gamma_1, ..., \Gamma_{p-1}$ *and* Ω*, the conditional estimator for* γ *under the restriction (20) is given by (18) with* H_γ *replaced with* H*.*

Proof. Similarly to the proof of Theorem 4. ■

3.1.3 Restrictions on β

Estimates on α and β are obtained by standard reduced rank regression. The only modification we apply here is that we condition on γ. Let $Z_{0t} = \Delta Y_t - \gamma \Delta D_t$, $Z_{1t} = vec\left(Y_{t-1} - \gamma D_{t-1}, 1\right)$ and $Z_{2t} = vec\left(\Delta Y_{t-1} - \gamma \Delta D_{t-1}, ..., \Delta Y_{t-p+1} - \gamma \Delta D_{t-p+1}\right)$. Furthermore, let R_{0t} and R_{1t} be the residuals obtained by regressing Z_{0t} and Z_{1t} on Z_{2t}, respectively. The system in (12) can then by written as

$$R_{0t} = \alpha \beta^{*\prime} R_{1t} + \varepsilon_t. \tag{21}$$

The estimator for α is

$$\widehat{\alpha}\left(\beta\right) = S_{01}\beta\left(\beta' S_{11} \beta\right)^{-1}, \tag{22}$$

where

$$S_{ij} = T^{-1} \sum_{t=1}^{T} R_{it} R_{jt}'.$$

The cointegration space is found by solving

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

for the eigenvalues $1 > \widehat{\lambda}_1 > ... > \widehat{\lambda}_n > 0$ and eigenvectors $\widehat{V} = (\widehat{v}_1, ..., \widehat{v}_n)$, which we normalise by $\widehat{V} S_{11} \widehat{V} = I$. Then, $\widehat{\beta} = (\widehat{v}_1, ..., \widehat{v}_r)$, see Johansen (1995, p.93).

If restrictions of the form $R_\beta' \beta = 0$ are imposed on the cointegration space, we solve the eigenvalue problem $\left| \lambda H_\beta' S_{11} H_\beta - H_\beta' S_{10} S_{00}^{-1} S_{01} H_\beta \right| = 0$ for $H_\beta = \left(R_\beta'\right)_\perp$ and finally $\beta = H_\beta \cdot (\widehat{v}_1, ..., \widehat{v}_r)$.

3.1.4 Switching algorithm

The estimators presented here are conditional estimators, i.e. they are conditional on other parameters in the model. They will therefore only reveal the full information maximum likelihood estimates if they condition on the full information maximum likelihood values of the conditional parameters. However, a switching algorithm may be used in order to obtain the maximum likelihood estimates. The full information maximum likelihood estimates may then be found by switching between estimating γ conditioned on the other parameters and the other parameters conditioned on γ until convergence.[4]

[4] However, this might be very time consuming. To speed up the algoritm, one could apply a simulation algorithm such as the quasi-Newton method developed by Broyden, Fletcher, Goldfarb and Shannon (BFGS). This maximation algoritm is implemented in Ox (MaxBFGS).

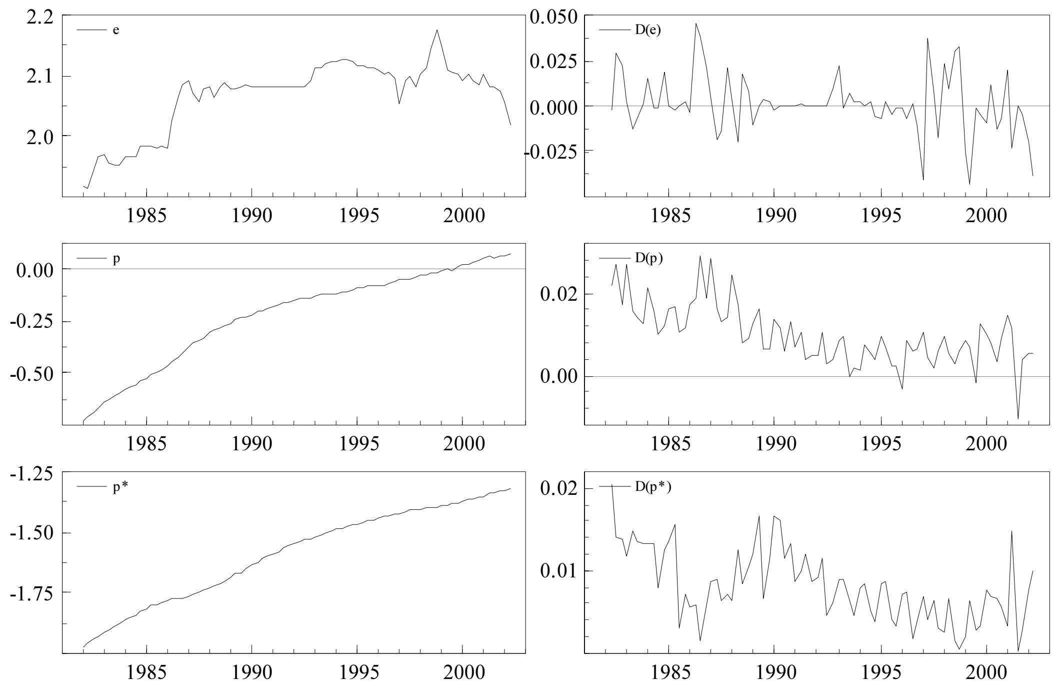

Figure 2: The exchange rate (e), Norwegian consumper prices (p) and foreign consumer prices (p^*) in the left column. The difference of the data in the right column.

4 Empirical illustrations

4.1 Purchasing power parity

To illustrate the estimation method we here analyse a data set with three variables; Norwegian and foreign prices (p and p^* respectively) and the corresponding exchange rate (e). The data, $Y_t' = (e_t, p_t, p_t^*)$, are plotted in Figure 2. The inflation rate was higher in the 1980's than in the 1990's. We therefore include a step dummy picking up this shift in inflation. The step dummy is (arbitrarily) chosen to be zero until the end of 1989 and unity from the first quarter of 1990, as in Figure 1.

We allow a broken trend in the data. Therefore, both the constant and the step dummy is included also accumulated. The vector of deterministic variables are therefore $D = (t, D1990q1_t, cum\,(D1990q1_t))$, where $cum\,(D1990q1_t)$ is the accumulated value of $D1990q1_t$. In addition, (centered) seasonal variables are included in the empirical analysis. The estimation period is 1983q1-2002q2.

The cointegration rank test is reported in Table 2. [No probability values are reported. These are not standard critical values, since they will depend on the break point. In the following we will assume one cointegrating vector.]

We first test if the real exchange rate is trend-stationary with a possible break in both level and trend from the first quarter of 1990. The test results are reported in Table 3. These restrictions on the cointegration vector are not rejected.

Table 2: Cointegration analysis

Rank:	$-T/2\log\lvert\Omega\rvert$	loglik	Hypothesis	Trace	p-value
$r = 3$	1251.54				
$r = 2$	1249.51		$r \leq 2$	4.05	
$r = 1$	1240.68		$r \leq 1$	21.7	
$r = 0$	1222.94		$r = 0$	57.2	

Table 3: Likelihood ratio test: Restrictions on the cointegration vector

Test	$-T/2\log\lvert\Omega\rvert$	loglik	LR	d.f.	p-value
No restrictions	1240.68				
Real exchange rate	1238.78		3.80	2	[0.15]

In Table 4 different types of test of structural breaks in the system are tested. In the upper part of the table the restrictions on the cointegration vector are not imposed. Therefore, the unrestricted model with one cointegrating vector is the alternative hypothesis.

In the lower part the restrictions on the cointegration vector are imposed. The results of the different tests of structural breaks are independent of whether the restrictions on the cointegration vector are imposed or not. Therefore we discuss the two parts of the table together.

First we test breaks in the cointegrated vector. Both the hypothesis of no break in trend and the hypothesis of no break in level are accepted (though the former test is close to be rejected at the 5 per cent level when the restrictions on the cointegration vector are imposed). The combined hypothesis is also accepted.

Testing for structural breaks in the levels the hypothesis of a level shift is accepted. (This hypothesis also implies no shift in the cointegration vector.) The hypothesis of no break in trends is rejected at the 5 per cent level when the cointegration vector is restriced, and close to be rejected at the same significance level when the cointegration vector is not restricted.

The last test is a combined test, where we test for no breaks in the cointegrating vector and no level shifts in the variables. This combined test is accepted.

Below the system is reported when the combined restrictions are imposed. Also the restrictions on the cointegration vector are imposed, and seasonality is not re-

Table 4: Likelihood ratio tests: Structural breaks in cointegration vector and variables

Test	$-T/2 \log \lvert\Omega\rvert$	loglik	LR	d.f.	p-value
No restrictions	1240.68				
No break in trend in CI-vector	1240.37		0.62	1	[0.43]
No break in level	1240.66		0.03	1	[0.86]
No break in level nor trend	1240.24		0.88	2	[0.65]
No break in trend in variables	1236.97		7.41	3	[0.06]
No break in level	1239.49		2.37	3	[0.50]
No break in level nor trend	1234.76		11.84	6	[0.07]
No break in CI-vector & no break in levels in variables	1239.08		3.19	4	[0.53]
Test	$-T/2 \log \lvert\Omega\rvert$	loglik	LR	d.f.	p-value
Real exchange rate	1238.78				
No break in trend in CI-vector	1237.06		3.44	1	[0.06]
No break in level	1238.77		0.02	1	[0.89]
No break in level nor trend	1237.05		3.45	2	[0.18]
No break in trend in variables	1233.42		10.72	3	[0.01]*
No break in level	1236.84		3.86	3	[0.28]
No break in level nor trend	1231.44		14.67	6	[0.02]*
No break in CI-vector & no break in levels in variables	1235.30		6.95	4	[0.14]

ported.

$$
\begin{aligned}
&\begin{pmatrix} \Delta e \\ \Delta p \\ \Delta p^* \end{pmatrix}_t - \begin{pmatrix} 0 & 0.0077 & -0.0106 \\ 0 & 0.0123 & -0.0066 \\ 0 & 0.0021 & 0.0040 \end{pmatrix} \begin{pmatrix} \Delta D1990q1 \\ 1 \\ D1990q1 \end{pmatrix}_t \\
= &\begin{pmatrix} -0.064 \\ -0.026 \\ -0.039 \end{pmatrix} \left\{ \begin{pmatrix} 1 & -1 & 1 \end{pmatrix} \right. \\
&\times \left. \left[\begin{pmatrix} e \\ p \\ p^* \end{pmatrix}_{t-1} - \begin{pmatrix} 0 & 0.0077 & -0.0106 \\ 0 & 0.0123 & -0.0066 \\ 0 & 0.0021 & 0.0040 \end{pmatrix} \begin{pmatrix} D1990q1 \\ t \\ cum(D1990q1) \end{pmatrix}_{t-1} \right] - 0.852 \right\} \\
&+\Gamma_1 \left[\begin{pmatrix} \Delta e \\ \Delta p \\ \Delta p^* \end{pmatrix}_{t-1} - \begin{pmatrix} 0 & 0.0077 & -0.0106 \\ 0 & 0.0123 & -0.0066 \\ 0 & 0.0021 & 0.0040 \end{pmatrix} \begin{pmatrix} \Delta D1990q1 \\ 1 \\ D1990q1 \end{pmatrix}_{t-1} \right] \\
&+seasonality + \varepsilon_t
\end{aligned}
$$

where

$$\begin{pmatrix} 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0.0077 & -0.0106 \\ 0 & 0.0123 & -0.0066 \\ 0 & 0.0021 & 0.0040 \end{pmatrix} \begin{pmatrix} D1990q1 \\ t \\ cum(D1990q1) \end{pmatrix}_{t-1}$$
$$= \begin{pmatrix} 0 & -0.0024 & 0 \end{pmatrix} \begin{pmatrix} D1990q1 \\ t \\ cum(D1990q1) \end{pmatrix}_{t-1}$$

[Standard deviations not yet calculated.]

From the estimation result a domestic quarterly underlying inflation of 1.23 per cent in the 1980's is identified. This corresponds to an annual inflation of about 5 per cent. From 1990 the quarterly equilibrium inflation rate is 0.5 per cent $(1.23 - 0.66)$, which corresponds to an annual inflation of 2.30 per cent.

4.2 Uncovered interest parity

Johansen et al. (2000) apply their method to analyse the uncovered interest parity (UIP) hypothesis between Germany and Italy. Here we use the same data with our method. The data used in the analysis are first differences of log consumer price indices for Italy and Germany $(\Delta p_t^I, \Delta p_t^D)$; the first difference of log nominal exchange rate between Italian Lira and German Mark (Δe_{t+1}) representing the rational expectation to future exchange rates; and nominal interest rates on long-term treasury bonds in both countries (i_t^I, i_t^D). The data, $Y_t' = \left(\Delta p_t^I, \Delta p_t^D, \Delta e_{t+1}, i_t^I, i_t^D\right)$, is plotted in the right column in Figure 3.[5] Johansen et al. (2000) introduce two breaks; 1980q1 and 1992q3. The former corresponds to the creation of the EMS (but is also supposed to catch the oil price shock and the modification on the US monetary policy), and the latter corresponds to the exit of Italy from the EMS and the reunification of Germany. The vector of deterministic variables is therefore given by $D_t = (t, D1980q1_t, D1992q3_t, cum\,(D1980q1_t)\,, cum\,(D1992q3_t))$.

In Figure 3 we see that there are significant trends in inflation and interest rates at least in some of the periods. Therefore, we use the model H_l. The estimation period is 1973q4-1995q4, and the number of lags is 2 ($p = 2$).

The cointegration rank test is reported in Table 5. [No probability values are reported. These are not standard critical values, since they will depend on the break point. In the following we will assume three cointegrating vectors.]

Following Johansen et al. (2000), we suggest the three stationary linear combinations reported below. They correspond to the UIP hypothesis, the German real interest rate, and the real interest rate differential.

$$\begin{aligned} y_{1t} &= i_t^I - \left(i_t^D + \Delta e_{t+1}\right) \\ y_{2t} &= \left(i_t^D - \Delta p_t^D\right) \\ y_{3t} &= \left(i_t^I - \Delta p_t^I\right) - \left(i_t^D - \Delta p_t^D\right) \end{aligned}$$

[5] The data are available at http://www.blackwellpublishers.co.uk/ectj/dataset5.htm. Source: Johansen et al. (2000).

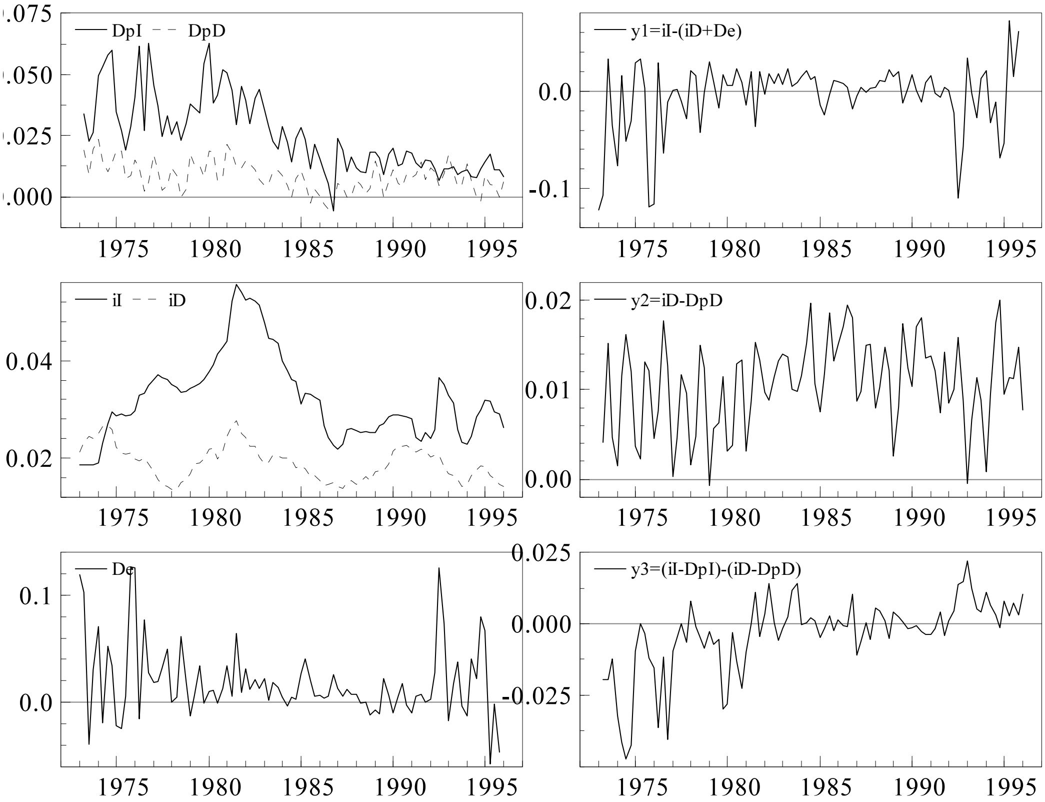

Figure 3: The data in the left column: inflation in Italy (DpI) and Germany (DpD); interest rates in Italy (iI) and Germany (iD); and log difference of LIT/DM exchange rate (De). In the rigth column are the three components we expect to be stationary.

Table 5: Cointegration rank test: ...

Rank:	loglik	Hypothesis	Trace	p-value
$r = 5$	1865.07			
$r = 4$	1860.48	$r \leq 4$	9.19	
$r = 3$	1852.97	$r \leq 3$	24.22	
$r = 2$	1832.76	$r \leq 2$	64.62	
$r = 1$	1792.99	$r \leq 1$	144.16	
$r = 0$	1748.71	$r = 0$	232.74	

Table 6: Likelihood ratio test: Restrictions on the cointegration space

Test	loglik	LR	d.f.	p-value
No restrictions	1852.97			
Restrictions on coint. space	1844.74	16.45	6	0.012

Table 7: Likelihood ratio tests: Structural breaks in cointegration vectors

Test	loglik	LR	d.f.	p-value
No restrictions	1852.97			
No trend in first period	1848.68	8.58	3	0.035
No trend in second period	1845.56	14.81	3	0.002
No trend in third period	1834.72	36.49	3	0.000
No trend-break from 1st to 2nd	1847.72	10.50	3	0.015
No trend-break from 2nd to 3rd	1834.74	36.46	3	0.000
No level-break from 1st to 2nd	1845.39	15.15	3	0.002
No level-break from 2nd to 3rd	1837.75	29.53	3	0.000
Test	loglik	LR	d.f.	p-value
Restricted cointegration space	1844.74			
No trend in first period	1838.15	13.19	3	0.004
No trend in second period	1843.33	2.83	3	0.419
No trend in third period	1826.34	36.81	3	0.000

These suggested stationary components imply that the cointegration space should span the space of the matrix reported below:

$$\beta' \in sp \begin{pmatrix} 0 & 0 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{pmatrix}$$

The restrictions imposed on the cointegration space augmented with an intercept (i.e. $\beta^{*\prime} = (\beta', -\mu)$) is therefore

$$R'_\beta = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \end{pmatrix},$$

where the left 2×5 matrix (which involve the restrictions on β) are orthogonal to the cointegration space, and the right vector with zeros corresponds to that we have not imposed any restrictions on the intercept in the cointegration space.

The restriction implied by the suggested cointegration space is not accepted at a 5 per cent level, see Table 6. In the following we test different types of breaks both with and without the suggested cointegrated space.

In the upper part of Table 7 different test of breaks in trends and levels in the cointegration space are tested when the restriction on the cointegration space is not imposed. Some of the tests are repeated in the bottom part of Table 7 when the cointegration restrictions are imposed.

The first three hypotheses in Table 7 corresponds to the hypotheses in Table 9 in Johansen et al. (2000). For each of the three periods the hypotheses of no trend is

Table 8: Likelihood ratio tests: Structural breaks in variables

Test	loglik	LR	d.f.	p-value
No restrictions	1852.97			
No trend in first period	1846.47	13.00	5	0.023
No trend in second period	1845.34	15.25	5	0.009
No trend in third period	1834.06	37.81	5	0.000
Test	loglik	LR	d.f.	p-value
Restricted coint. space	1844.74			
No trend in first period	1836.72	16.04	5	0.007
No trend in second period	1843.12	3.248	5	0.662
No trend in third period	1825.79	37.91	5	0.000

rejected at a 5 per cent level. The reported probability value is somewhat smaller than the corresponding probability values reported in Johansen et al. (2000). The discrepancies steams form the fact that they include impulse dummies in the two periods after the break.[6]

In the next two lines of Table 7 hypotheses of no break between the different periods are tested. Both hypotheses are rejected. These tests could also be tested by the method in Johansen et al. (2000).

However, tests of level-breaks can not be tested by the method suggested in Johansen et al. (2000). The test results reported in Table 7 show that such hypotheses are rejected.

The three first restrictions are re-tested when the restrictions on the cointegration space are imposed.[7] Also now the restrictions that there are no trends in the cointegration space are rejected for the first and third period. In the second period, however, the restriction is now far from being rejected. This can also be seen from the right column of Figure 3: In the second period (1980q1-1992q2) there are no trends in the suggested stationary relationships.

Also tests on the trends in the variables can be conducted. In Table 8 these tests are reported. Again we test both with and without the restrictions imposed by the suggested cointegration space.

From the upper part of Table 8 we see that the hypotheses that there is no trend in any of the periods can be rejected when no restrictions are imposed on the cointegration space. When restrictions are imposed on the cointegration space, however, we can not reject the hypothesis that there is no trend in the data in the second period.

[6] In our framework we can get test results that are approximatly equal to Johansen et al. (2000) by extending the vector of deterministic variables with $(D1980q2_t, D1980q3_t, D1992q4_t, D1993q1_t)$.

[7] We analyse the system also with these restrictions on the cointegrating space imposed even though they were rejected (see Table 6), because they were included in the analysis in Johansen et al. (2000). They found that these restrictions in the could be accepted when combined with additional restrictions on the trends in the cointegrating space.

5 Conclusions

In this paper we have showed how to decompose deterministic parts into interpretable long run counterparts. The estimation procedure also allows for conducting hypothesis testing on the corresponding coefficients. We illustrate one application of the procedure by analysing data sets with structural breaks, and show how to test for different types of structural breaks.

6 References

Boswijk, P. (1995). Identifibility of Cointegrated Systems. Discussion Paper ti 7-95-078, Tinbergen Institute, University of Amsterdam.

Hungnes, H. (2002). Restricting Growth Rates in Cointegrated VAR Models. Revised version of Discussion Papers 309, Statistics Norway. (Downloadable at http://home.chello.no/~hhungnes/Publications/DP309rev.pdf)

Johansen, S., R. Mosconi and B. Nielsen (2000). Cointegration Analysis in the Presence of Structural Breaks in the Deterministic Trend. Econometrics Journal, 3:216-249.

Lütkepohl, H. (1996). Handbook of Matrices. New York: John Wiley & Sons.

Saikkonen, P. and H. Lütkepohl (2000). Testing for the Cointegrating Rank of a VAR Process with Structural Shifts. Journal of Business & Economic Statistics, 18:451-464.

Magnus, J. R. and H. Neudecker (1979). Matrix Differential Calculus with Applications in Statistics and Econometrics. New York: John Wiley & Sons.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57:1361-1401. Erratum (1993) Econometrica 61:248-249.

Theil, H. (1983). Linear Algebra and Matrix Methods in Econometrics. In Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics, Vol 1.* New Yory: North Holland.

7 Appendix

7.1 Proof of Theorem 4

Proof. In the proof of the theorem we use $tr\,(AB) = tr\,(BA) = (vecA')'\,vecB$ and $vec\,(AXB) = (B' \otimes A)\,vecX$ where tr is the trace operator. The log-likelihood function can be written as

$$\log L = -\frac{T}{2}\left[\begin{array}{c} tr\Omega^{-1}\Phi S_{ZZ}\Phi' + tr\Omega^{-1}\alpha\beta^{*\prime}S_{YY}\beta^{*}\alpha' \\ -tr\Omega^{-1}\Phi S_{ZY}\beta^{*}\alpha' - tr\Omega^{-1}\alpha\beta^{*\prime}S_{YZ}\Phi' \\ +N + tr\Omega^{-1}\Phi^{*}\left(I_{p+1}\otimes\gamma\right)S_{DD}\left(I_{p+1}\otimes\gamma'\right)\Phi^{*\prime} + N' \end{array}\right], \tag{23}$$

where

$$\begin{aligned} N &= -tr\Omega^{-1}\Phi S_{ZD}\left(I_{p+1}\otimes\gamma'\right)\Phi^{*\prime} + tr\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\left(I_{p+1}\otimes\gamma'\right)\Phi^{*\prime} \\ &= -tr\left(I_{p+1}\otimes\gamma'\right)\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\ &= -\left(vec\left(I_{p+1}\otimes\gamma\right)\right)'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\ &= -\left(vec\gamma\right)'M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\ &= -\psi' H_{\gamma}' M' vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \end{aligned}$$

The derivative of (23) with respect to ψ (under the restrictions in (14)) is

$$\begin{aligned} &\frac{\partial \log L}{\partial \psi} \qquad\qquad (24) \\ &= TH_{\gamma}'M'\left[vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD} - \Phi^{*\prime}\Omega^{-1}\Phi^{*}\left(I_{p+1}\otimes\gamma\right)S_{DD}\right)\right]. \end{aligned}$$

Setting (24) equal to zero and using (14) and (17) lead to

$$\begin{aligned} 0 &= H_{\gamma}'M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\ &\quad -H_{\gamma}'M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi^{*}\left(I_p\otimes\gamma\right)S_{DD}\right) \\ &= H_{\gamma}'M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\ &\quad -H_{\gamma}'M'\left(S_{DD}\otimes\Phi^{*\prime}\Omega^{-1}\Phi^{*}\right)vec\left(I_p\otimes\gamma\right) \\ &= H_{\gamma}'M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\ &\quad -H_{\gamma}'M'\left(S_{DD}\otimes\Phi^{*\prime}\Omega^{-1}\Phi^{*}\right)MH_{\gamma}'\psi \\ &= H_{\gamma}'M'\left[vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right)\right] \\ &\quad -\left[H_{\gamma}'M'\left(S_{DD}\otimes\Phi^{*\prime}\Omega^{-1}\Phi^{*}\right)MH_{\gamma}'\right]\psi \end{aligned}$$

Solving for ψ and using (14) again lead to (18). ∎