Central European Journal of Economic Modelling and Econometrics

Common Trends and Common Cycles – Bayesian Approach

Justyna Wróblewska*

Submitted: 10.10.2014, Accepted: 21.06.2015

Abstract

In 1993 Engle and Kozicki proposed the notion of common features of which one example is a serial correlation common feature. We say that stationary, non-innovation processes exhibit common serial correlation when there exists at least one linear combination of them which is an innovation. Later on in 1993 Vahid and Engle combined the notions of cointegration among I(1) processes with common serial correlation within their first differences. It is commonly known that cointegrated time series have vector error correction (VEC) representation. The existence of common serial correlation leads to an additional reduced rank restriction imposed on the VEC model's parameters. This type of restriction was later termed a strong form (SF) reduced rank structure, as opposed to a weak one introduced in 2006 by Hecq, Palm and Urbain.
The main aim of the present paper is to construct the Bayesian vector error correction model with these additional strong form restrictions.
The empirical validity of investigating both the short- and long-run co-movements between macroeconomic time series will be illustrated by the analysis of the price-wage nexus in the Polish economy.

Keywords: cointegration, Bayesian analysis, common cyclical features, matrix Bingham-von Mises-Fisher distribution, matrix Langevin-Bingham distribution

JEL Classification: C11, C15, C32

*Cracow University of Economics; e-mail: eowroble@cyf-kr.edu.pl

1 Introduction

While analysing multivariate time series there is of natural interest to check whether the examined variables move together in the long and short run. There is a wide literature explaining and applying the concept of cointegration, which is an indicator of long-run co-movement between integrated series. The integrated series are cointegrated when there exists at least one non-zero linear combination of them which lowers the integration rank. Engle and Kozicki (1993) proposed the notion of common features, which can be regarded as the extension of the concept of cointegration. They considered *features* satisfying three axioms: *i.* if x has (does not have) the feature, then λx ($\lambda \neq 0$) will have (not have) it; *ii.* if both series x and y do not have the feature, then the sum $x+y$ will not have it either; *iii.* if only one of the series x or y has the feature, then the sum $x+y$ will also have it. According to the definition they have formulated, *a feature that is present in each of a group of series is said to be common to those series if there exists a non-zero linear combination of the series that does not have the feature* (Engle, Kozicki 1993, p. 370). One of the example of the common feature idea is a common serial correlation (or *a non-innovation co-feature* as proposed by Ericsson 1993). We say that stationary, non-innovation processes exhibit a common serial correlation when there exists at least one linear combination of them which is an innovation.

Vahid and Engle (1993) combined the notions of cointegration among processes integrated of order 1 (I(1)) with common serial correlation among their first differences, so it is possible to analyse the short- and long-run co-movement of the series in one model.

It is commonly known that cointegrated time series have vector error correction (VEC) representation. Vahid and Engle (1993) showed that the existence of a common serial correlation leads to an additional reduced rank restriction imposed on the VEC model parameters. Hecq, Palm and Urbain (2006) termed this type of commonality the strong form reduced rank structures, in opposition to weak and mixed forms introduced by them. When the first differences of CI(1,1) series (i.e. I(1) processes which cointegrate to I(0)) exhibit the strong form co-dependence, the adjustment coefficients matrix and the matrices describing the behaviour of the first differences have a common left null space, whereas in the case of the weak form only the left null spaces of matrices for the first differences have to overlap. The mixed form combines both of the previously mentioned, i.e. in the set of co-feature vectors there is one group which fulfils the assumptions for the weak form and the second, which also satisfies conditions for the strong form. All of these types differ also in terms of interpretation (see Hecq, Palm and Urbain (2006) for details).

The main aim of the paper is to construct the Bayesian vector error correction model with additional strong form restrictions (Bayesian VEC-SF model). The idea of how to unambiguously identify the model parameters comes from the papers by Strachan and Inder (2004), Koop, León-González and Strachan (2010), and Villani (2005), so the identification will be achieved *via* imposing orthonormal restrictions on two sets of

vectors: the one describing the long-run behaviour and the second one - representing the short-run dependence.
The empirical validity of investigating both the short- and long-run co-movements of macroeconomic time series will be illustrated by the analysis of the price-wage nexus in the Polish economy. The research will be based on 72 quarterly observations of 5 economic categories (nominal wages, consumer price index, productivity, rate of unemployment, import prices) covering the period 1995-2012.

2 The model

Let us write the n-dimensional cointegrated process in the VEC form:

$$\Delta x_t = \alpha\beta'\tilde{x}_{t-1} + \sum_{i=1}^{k-1}\Gamma_i\Delta x_{t-i} + \Phi D_t + \varepsilon_t = \alpha\beta'\tilde{x}_{t-1} + \Gamma' z_t + \Phi D_t + \varepsilon_t, \qquad (1)$$

where $\tilde{x}_{t,\ m\times 1} = (x_t',\ d_t)'$, $z_t' = (\Delta x_{t-1}', \Delta x_{t-2}', \ldots, \Delta x_{t-k+1}')$, $\Gamma = (\Gamma_1, \ldots, \Gamma_{k-1})'$, $\varepsilon_t \sim iiN^n(0, \Sigma)$, $t = 1, 2, \ldots, T$, D_t and d_t contain deterministic components. The matrices $\beta_{m\times r}$ and $\alpha_{n\times r}$ have full column rank, $m = n$ when no deterministic components are included in the cointegrating relations, otherwise $m > n$ and $\beta' = (\beta^{(1)'},\ \beta^{(2)'})'$, where $\beta^{(1)}$ is an $n \times r$ matrix and $\beta^{(2)}$ - $(m-n)\times r$ one.
It is commonly known that the VEC model suffers from non-identification, so to unambiguously estimate matrices β and α one has to imposed a total of r^2 restrictions on these parameters. Most often all the restrictions are put on the cointegration matrix β. In other words we can say that the cointegration space is the only thing one could get information from the data about. Following the idea by Strachan and Inder (2004), and Villani (2005), we assume that β has orthonormal columns, so $\beta'\beta = I_r$. Additionally we assume that first elements in each column of β are positive. This last restriction help us to deal with the many-to-one relation between matrices with orthonormal columns and the space they span.
The cointegrated VAR model allows for modelling of both the long-run relationships between time series and the short-run dynamics.
The cointegration is an evidence of co-movement among non-stationary time series. According to the idea of co-dependence, stationary time series can also move together. In such a case there exists at least one linear combination of them which lowers the moving average order. As pointed out by Vahid and Engle (1993), the serial correlation common feature introduced by Engle and Kozicki (1993) is a strong form of co-dependence. Non-innovation time series share serial correlation common feature if there exists at least one linear combination of them which is an innovation. In 1993, Vahid and Engle proposed the model combining both ideas - the cointegration and the serial correlation common feature.
The existence of serial correlation common features leads to additional rank restrictions in the VEC model and in the multivariate Beveridge-Nelson trend-cycle

decomposition of cointegrated series, which reads as follows:

$$x_t = \delta_t + \tau_t + \kappa_t = \delta_t + C(1)\sum_{i=0}^{t-1}\varepsilon_{t-i} + C^*(L)\varepsilon_t, \tag{2}$$

where δ_t denotes the deterministic component, τ_t - the trend, κ_t- the stationary part, $C(1) = \beta_{\perp}^{(1)}(\alpha'_{\perp}(I_n - \sum_{i=1}^{k-1}\Gamma_i)\beta_{\perp}^{(1)})^{-1}\alpha'_{\perp}$, $C^*(L)\varepsilon_t = \sum_{j=0}^{\infty}C_j^*\varepsilon_{t-j}$, $C_j^* = -\sum_{i>j}C_i$ and $C_i \quad (i = 0, 1, \dots)$ are matrices from the Wold representation of Δx_t, which by our assumptions is I(0), so they fulfil the inequality $\sum_{i=1}^{\infty} i|C_i| < \infty$. For any $(m \times s)$ full column rank matrix C, $C_\perp$ denotes an $(m \times (m-s))$ matrix of full column rank such that $C'C_\perp = 0$.
In an n-dimensional CI(1,1) process with r cointegrating vectors the matrix $C(1)$ is of a reduced rank equal to $n-r$, so the stochastic trend of the process is defined by $n-r$ random walks. Vahid and Engle (1993) asked similar question about the stationary (cyclical) part of the process and they proved that it is possible to eliminate the cyclical part by linear combination when all the matrices $C_j^*,\ j = 0, 1, \dots$ are of reduced-rank and their left null spaces overlap. This assumption is equivalent to reduced-rank restrictions for the matrices $C_i,\ i = 1, 2, \dots,$ from the Wald decomposition of the first differences of the analysed process, so *the same transformation that eliminates serial correlation in the differences when applied to the levels eliminates the common cycles* (Vahid and Engle 1993, p. 344). They also point out that the number of linearly independent vectors which eliminate the common cycles cannot exceed the number of common trends $(n-r)$ and these vectors are linearly independent of the cointegrating vectors.
Additionally, Vahid and Engle (1993) demonstrate consequences of common serial correlation for the VEC representation of cointegrated series. As it was previously recalled, in the case of the strong co-dependence of Δx_t there exists at least one linear combination of its components, which does not depend on the past values, so there should exist at least one vector $\gamma_\perp^*$ such that $\gamma_\perp^{*'}(\Delta x_t - \Phi D_t) = \gamma_\perp^{*'}\varepsilon_t$. The vectors collected in $\gamma_\perp^*$ are called co-feature vectors, and the combinations $\gamma_\perp^{*'}(\Delta x_t - \Phi D_t)$ are called co-feature combinations.
The common serial correlation assumption leads to an additional rank reduction in the VEC model as $\gamma_\perp^{*'}(\alpha\beta'\tilde{x}_{t-1} + \Gamma' z_t)$ should equal zero and this equality implies that α and all Γ's have a lower than the full rank and their left null spaces must overlap, so they possess the following representations: $\alpha = \gamma^*\delta_0^{*'},\ \Gamma_i = \gamma^*\delta_i^{*'},\ i = 1, 2, \dots, k-1$, where the matrix γ^* is of full column rank equal to $n-s$, where s denotes the number of linearly independent co-feature vectors.
Such factorisations are, of course, not unique, i.e. for any non-singular matrix (say M_i) of appropriate dimension the following equalities hold: $\gamma^* M_i M_i^{-1}\delta_i^{*'} = \gamma^*\delta_i^{*'},\ i = 0, 1, \dots, k-1,$ so in order to unambiguously estimate the model parameters we have to impose additional identification restrictions on them. We will use the method proposed by Koop, León-González and Strachan (2010). Before explaining the idea of this technique let us write the model combining cointegration

and common serial correlation features:

$$\Delta x_t = \gamma^* \delta_0^{*'} \beta' \tilde{x}_{t-1} + \sum_{i=1}^{k-1} \gamma^* \delta_i^{*'} \Delta x_{t-i} + \Phi D_t + \varepsilon_t = \gamma^* \delta^{*'} z_t^* + \Phi D_t + \varepsilon_t, \quad \varepsilon_t \sim iiN^n(0, \Sigma) \tag{3}$$

where $\delta^{*'} = (\delta_0^{*'}, \delta_1^{*'}, \ldots, \delta_{k-1}^{*'})$, $z_t^* = (\tilde{x}'_{t-1}\beta, \Delta x'_{t-1}, \ldots, \Delta x'_{t-k+1})'$, Σ is a positive-definite and symmetric matrix, the initial conditions $x_{-k+1}, \ldots, x_0$ are treated as known. The matrices $\beta_{m\times r}$ and $\gamma^*_{n\times(n-s)}$ are of full column rank.
Hecq, Palm and Urbain (2006) termed this type of commonality as the strong form reduced rank structures, in opposition to the weak and mixed forms introduced by them. We will use their notation and label model (3) as the VEC-SF model. To save space let us represent the VEC-SF model in matrix notation:

$$Z_0 = \begin{pmatrix} Z_1\beta, & Z_2 \end{pmatrix} \delta^* \gamma^{*'} + Z_3 \Gamma_s + E, \tag{4}$$

where

$$Z_0 = (\Delta x_1,\ \Delta x_t,\ \ldots,\ \Delta x_T)',$$

$$Z_1 = (\tilde{x}_0,\ \tilde{x}_1,\ \ldots,\ \tilde{x}_{T-1})',$$

$$Z_2 = (z'_1,\ z'_2,\ \ldots,\ z'_T)' \text{ with } z_t = \left(\Delta x'_{t-1},\ \Delta x'_{t-2},\ \ldots,\ \Delta x'_{t-k+1}\right),$$

$$Z_3 = (D_1,\ D_2,\ \ldots,\ D_T)'$$

and $\Gamma_s = \Phi'$.
To complete the definition of the Bayesian VEC-SF model one has to settle the prior structure. But before that it is necessary to establish the identification restrictions imposed on γ^* and δ^*. Following the concept by Koop, León-González and Strachan (2010) we will consider two equivalent parameterisations of the reduced rank matrix $\begin{pmatrix} \alpha, & \Gamma' \end{pmatrix}'$:

$$D^* G^{*'} = \delta^* \gamma^{*'}.$$

In the first parameterisation $D^*_{(n(k-1)+r)\times q} \in \mathbb{R}^{(n(k-1)+r)q}$, $G^*_{n\times q} \in \mathbb{R}^{nq}$, $q = n - s$ and $r(G^*) = q$, whereas in the second one we assume that γ^* has orthonormal columns and, additionally, that γ^* has positive elements in the first row, i.e. $\gamma^* \in \tilde{V}_{q,n}$, which is the 2^{-q}th part of the Stiefel manifold containing $n \times q$ matrices with orthonormal columns (see Chikuse 2002 for a general description). To identify β, which is the matrix containing vectors spanning the cointegration space, we assume that $\beta'\beta = I_r$ and that it has positive elements in the first row, so $\beta \in \tilde{V}_{r,m}$.
We impose the following priors of the model parameters:

- a Matrix Angular Central Gaussian Distribution of β (truncated to $\tilde{V}_{r,m}$): $\beta | P_\tau, r, m \sim MACG(P_\tau)|_{\tilde{V}_{r,m}}$. Through P_τ the researcher may incorporate prior knowledge about the cointegration space. If we define $P_\tau = H_B H'_B +$

$+\tau H_{B\perp}H'_{B\perp}$, where $\tau \in [0,1]$, $H'_B H_B = I_r$, $H'_{B\perp}H_{B\perp} = I_{m-r}$, $H'_{B\perp}H_B = 0$, the scalar τ controls the tightness of the prior around the space spanned by the matrix H_B ($sp(H_B)$). The higher the scalar τ the less informative the prior. Imposing $\tau = 0$ one assumes that the cointegration space is $sp(H_B)$. Assuming $\tau = 1$ one obtains $P_\tau = I_m$, so a non-informative distribution of the cointegration space (see Chikuse 2002, Koop, León-González and Strachan 2010, and Strachan, Inder 2004, for the details);

- a matrix Normal of G^*, which leads to a Matrix Angular Central Gaussian Distribution of γ^* (truncated to $\tilde{V}_{q,n}$): $G^*|P_{\tau^*}, q \sim mN(0, \frac{1}{n}I_q, P_{\tau*})$, so $\tilde{\gamma}^* = G^*(G^{*'}G^*)^{-\frac{1}{2}} \sim MACG(P_{\tau^*})$ and, as we have assumed that γ^* has positive elements in the first row, we truncate this distribution to $\tilde{V}_{q,n}$ ($\gamma^* \sim MACG(P_{\tau^*})|_{\tilde{V}_{q,n}}$). Similarly to the above-stated distribution, prior knowledge is incorporated *via* the matrix $P_{\tau^*} = H_{G^*}H'_{G^*} + \tau^* H_{G^*\perp}H'_{G^*\perp}$, where $\tau^* \in [0,1]$, $H'_{G^*}H_{G^*} = I_q$, $H'_{G^*\perp}H_{G^*\perp} = I_{n-q}$, $H'_{G^*\perp}H_{G^*} = 0$;

- a matrix Normal of D^*: $D^*|k, r \sim mN(\underline{\mu}_{D^*}, \underline{\Omega}_{D^*}, \nu I_{n(k-1)+r})$, where $\underline{\Omega}_{D^*}$ is a positive-definite and symmetric matrix, ν is a positive constant, which can be estimated or set arbitrarily by a researcher;
  The assumption that γ^* is of full column rank is exploited in the construction of the Bayesian VEC-SF model and will be employed in the calculations of the posterior model probability. Note that, by imposing absolutely continuous distributions of D^* and G^* (so also of δ^*), we assume that, with probability 1, the matrix δ^* is of full column rank. Vahid and Engle (1993) proved that γ^* is of full column rank and the matrices δ_i^*, $i = 0, 1, \ldots, k-1$, need not be, admittedly, but note that δ^* groups all the matrices δ_i^*. A detailed reasoning concerning the rank of δ^* is a subject of a current research and is not presented here. Remain that imposing absolutely continuous (with respect to Lebesgue measure) prior distributions of D^* and G^* practically means that we work under the assumption of full column rank of δ^*.

- a matrix Normal of Γ_s: $\Gamma_s|h \sim mN(\underline{\mu}_s, \Sigma, hI_{l_s})$, where h is a positive constant, which can be estimated or set by a researcher, and l_s denotes the number columns of Γ_s;

- an inverted Wishart distribution of Σ: $\Sigma \sim iW(S, q_\Sigma)$, where S is a positive definite symmetric matrix;

  an inverted Gamma distribution of ν and h (if they are estimated):
  $h \sim iG(s_h, n_h)$, $\nu \sim iG(s_\nu, n_\nu)$; It follows that

$$E(h) = \frac{s_h}{n_h - 1} \quad \text{for} \quad n_h > 1$$

$$V(h) = \frac{s_h^2}{(n_h-1)^2(n_h-2)} \quad \text{for} \quad n_h > 2$$

and

$$E(\nu) = \frac{s_\nu}{n_\nu - 1} \quad \text{for} \quad n_\nu > 1$$

$$V(\nu) = \frac{s_\nu^2}{(n_\nu - 1)^2(n_\nu - 2)} \quad \text{for} \quad n_\nu > 2$$

- the constants τ and τ^*, which measure the informativeness of the MACG distributions, may be elicited by the researcher or estimated; if they are to be estimated, their prior densities should allocate most weight to values near zero and be restricted to [0, 1] (see Strachan and Inder 2004, and Koop, León-González and Strachan 2010); these could be e.g. inverted gammas truncated to [0, 1]: $\tau \sim iG_{[0,\ 1]}(s_\tau, n_\tau),\ \tau^* \sim iG_{[0,\ 1]}(s_{\tau^*}, n_{\tau^*})$.

The imposed priors lead to the following full conditional posteriors of the model parameters (see Appendix A for the calculations):

the full conditional posterior of Σ is an inverted Wishart distribution:

$$\Sigma|D^*, G^*, \Gamma_s, \beta, q, h, X \sim iW\left(S + \frac{1}{h}\left(\Gamma_s - \underline{\mu}_s\right)'\left(\Gamma_s - \underline{\mu}_s\right) + E'E, q_\Sigma + l_s + T\right)$$

where $E = Z_0 - \ (Z_1\beta, \quad Z_2)\ D^* G^{*'} - Z_3\Gamma_s$;

the full conditional posterior of $vec(D^*)$ is Normal:

$$vec(D^*)|\Sigma, G^*, \Gamma_s, \beta, q, h, \nu, X \sim N\left(\overline{\mu}_{vD^*}, \overline{\Omega}_{vD^*}\right),$$

where

$$\overline{\Omega}_{vD^*} = \left[\left(G^{*'}\Sigma^{-1}G^* \otimes \tilde{Z}'\tilde{Z}\right) + \left(\underline{\Omega}_{D^*}^{-1} \otimes \frac{1}{\nu} I_{n(k-1)+r}\right)\right]^{-1},$$

$$\overline{\mu}_{vD^*} = \overline{\Omega}_{vD^*} vec\left[\tilde{Z}'\left(Z_0 - Z_3\Gamma_s\right)\Sigma^{-1}G^* + \frac{1}{\nu}\underline{\mu}_{D^*}\underline{\Omega}_{D^*}^{-1}\right],$$

$\tilde{Z} = (Z_1\beta,\ Z_2)$ and $vec(F)$ denotes the vectorisation of F;

the full conditional posterior of $vec(G^*)$ is Normal:

$$vec(G^*)|\Sigma, D^*, \Gamma_s, \beta, q, h, \nu, X \sim N\left(\overline{\mu}_{vG^*}, \overline{\Omega}_{vG^*}\right),$$

where

$$\overline{\Omega}_{vG^*} = [(nI_q \otimes P_{\tau^*}^{-1}) + (D^{*'}\tilde{Z}'\tilde{Z}D^* \otimes \Sigma^{-1})]^{-1},$$

$$\mu_{vG^*} = \overline{\Omega}_{vG^*} vec[\Sigma^{-1}(Z_0 - Z_3\Gamma_s)'\tilde{Z}D^*];$$

the full conditional posterior of Γ_s is a matrix Normal:

$$\Gamma_s|\Sigma, G^*, D^*\beta, q, h, \nu, X \sim mN\left(\overline{\mu}_{\Gamma_s}, \Sigma, \left(\frac{1}{h}I_{l_s} + Z_3'Z_3\right)^{-1}\right),$$

where

$$\overline{\mu}_{\Gamma_s} = \left(\frac{1}{h}I_{l_s} + Z_3'Z_3\right)^{-1}\left[Z_3'\left(Z_0 - \tilde{Z}D^*G^{*'}\right) + \frac{1}{h}\underline{\mu}_s\right];$$

if $P_\tau = I_m$, then the full conditional posterior of β is a matrix Bingham-von Mises-Fisher distribution:

$$p(\beta|\Sigma, G^*, D^*\Gamma_s, r, q, h, \nu, X) \propto \exp\left\{tr\left(F'\beta + \tilde{B}\beta'\tilde{A}\beta\right)\right\}[\mathrm{d}\beta],$$

where $\tilde{A} = Z_1'Z_1$, $\tilde{B} = -\frac{1}{2}\tilde{\Gamma}\Sigma^{-1}\tilde{\Gamma}'$, $F = Z_1'\tilde{Z}\Sigma^{-1}\tilde{\Gamma}'$, $\tilde{Z} = Z_0 - Z_2\overline{D^*G^{*'}}$, $\overline{D^*G^{*'}}$ denotes the last $n(k-1)$ rows of the matrix $D^*G^{*'}$ and $\tilde{\Gamma}$ - the first r rows of $D^*G^{*'}$. The symbol $[\mathrm{d}\beta]$ denotes the normalised invariant measure on the Stiefel manifold.
This type of distribution is very flexible as it contains both linear and quadratic terms. It was introduced by Khatri and Mardia (1997). It combines the von Mises-Fisher distribution (also known as the matrix Langevin distribution, the linear term) and the Bingham distribution (the quadratic term), so it is called the matrix Bingham-von Mises-Fisher distribution or the matrix Langevin-Bingham distribution. The normalising constant for this distribution was given by De Waal (1979).
In the case of $P_\tau \neq I_m$, the full conditional posterior of β does not belong to any known class of distributions, but its density is proportional to the one of the matrix Bingham-von Mises-Fisher:

$$p(\beta|\Sigma, G^*, D^*\Gamma_s, q, h, \nu, X) \propto |\beta' P_\tau^{-1}\beta|^{-\frac{m}{2}} \exp\left\{tr\left(F'\beta + \tilde{B}\beta'\tilde{A}\beta\right)\right\}[\mathrm{d}\beta];$$

the full conditional posteriors of ν and h are inverted Gamma distributions:

$$\nu|\Sigma, G^*, q, X \sim iG\left(s_\nu + \frac{tr\left[\underline{\Omega}_{D^*}^{-1}\left(D^* - \underline{D^*}\right)'\left(D^* - \underline{D^*}\right)\right]}{2}, n_\nu + \frac{q[n(k-1)+r]}{2}\right)$$

$$h|\Sigma, \Gamma_s, X \sim iG(s_h + \frac{1}{2}tr[\Sigma^{-1}(\Gamma_s - \underline{\mu}_s)'(\Gamma_s - \underline{\mu}_s)], n_h + \frac{nl_s}{2});$$

the full conditional posterior of τ^* is an inverted Gamma truncated to $[0,1]$:

$$\tau^*|G^*, q, r, X \sim iG_{[0,1]}\left(s_{\tau^*} + \frac{1}{2}tr\left(nG^{*'}H_{G^*}^{\perp}H_{G^*}^{\perp'}G^*\right), n_{\tau^*} + \frac{1}{2}q(n-q)\right);$$

the full conditional posteriors of τ is of a non-standard form:

$$p(\tau|\beta, r, X) \propto |\beta' P_{\frac{1}{\tau}} \beta|^{-\frac{m}{2}} \exp\left(-\frac{s_\tau}{\tau}\right) \tau^{-n_\tau - \frac{1}{2}r(m-r)-1} \mathbf{1}_{[0,\ 1]}(\tau),$$

where $\mathbf{1}_{[0,\ 1]}(a)$ denotes a function which takes value of 1 for a in [0, 1] and 0 elsewhere.

3 Gibbs sampling algorithm

As the set of full conditional posterior distributions is known, the posterior results could be obtained by employing the Gibbs sampler. After choosing initial values ($\Sigma^{(0)}$, $D^{*(0)}$, $G^{*(0)}$, $\Gamma_s^{(0)}$, $\beta^{(0)}$, and, optionally, $\nu^{(0)}$, $h^{(0)}$, $\tau^{(0)}$, $\tau^{*(0)}$), the algorithm repeats the following steps:

Draw Σ from $iW\left(S + \frac{1}{h}\left(\Gamma_s - \underline{\mu}_s\right)'\left(\Gamma_s - \underline{\mu}_s\right) + E'E, q_\Sigma + l_s + T\right)$.

To obtain draws from the posterior distributions of δ^* and γ^* apply the method proposed by Koop, León-González and Strachan (2010) for VEC models, with a slight modification to take account of the additional restriction assuming positive elements in the first row of γ^*:

draw $vec(D^*)$ from $N\left(\overline{\mu}_{vD^*}, \overline{\Omega}_{vD^*}\right)$,

draw $vec(G^*)$ from $N\left(\overline{\mu}_{vG^*}, \overline{\Omega}_{vG^*}\right)$,

use the transformations:

$\delta^* = D^*(G^{*'}G^*)^{\frac{1}{2}}O_P$, $\gamma^* = G^*(G^{*'}G^*)^{-\frac{1}{2}}O_P$, where O_P denotes a diagonal matrix with 1 or -1 on its main diagonal, i.e. $O_P = diag(\pm 1)$.

Draw Γ_s from $mN\left(\overline{\mu}_s, \Sigma, \left(\frac{1}{h}I_{l_s} + Z_3'Z_3\right)^{-1}\right)$.

Draw β from $mBMF(\tilde{A}, \tilde{B}, F)$ in the case of the non-informative prior of cointegrating space (i.e. $P_\tau = I_m$).
One can use e.g. the Geodesic Monte Carlo on the Stiefel manifold constructed by Byrne and Girolami (2013). Another possibility is the Gibbs sampler proposed by Hoff (2009).
If $P_\tau \neq I_m$, one can use the Metropolis-Hastings algorithm with candidate values drawn from $mBMF(\tilde{A}, \tilde{B}, F)$, as in this case the posterior full conditional density of β is proportional the density of the distribution $mBMF(\tilde{A}, \tilde{B}, F)$, with the proportionality factor $|\beta' P_{\frac{1}{\tau}} \beta|^{-\frac{m}{2}}$.

Draw ν and h from the inverted Gamma distributions

$$iG\left(s_\nu + \frac{1}{2}tr\left[\underline{\Omega}_{D^*}^{-1}\left(D^* - \underline{D^*}\right)'\left(D^* - \underline{D^*}\right)\right], n_\nu + \frac{q[n(k-1)+r]}{2}\right)$$

and

$$iG\left(s_h + \frac{1}{2}tr\left[\Sigma^{-1}\left(\Gamma_s - \underline{\mu}_s\right)'\left(\Gamma_s - \underline{\mu}_s\right)\right], n_h + \frac{nl_s}{2}\right),$$

respectively.

Draw τ^* from the truncated inverted Gamma distribution

$$iG_{[0,1]}\left(s_{\tau^*} + \frac{1}{2}tr\left(nG^{*'}H_{G^*}^{\perp}H_{G^*}^{\perp'}G^*\right), n_{\tau^*} + \frac{1}{2}q(n-q)\right).$$

Draws from the non-standard distribution of τ could be obtained with the help of the Metropolis-Hastings algorithm with candidate values e.g. from $iG_{[0,\ 1]}\left(s_\tau, n_\tau + \frac{1}{2}r(m-r)\right)$, as in this case the posterior full conditional density of τ differs from the kernel of the density of the distribution $iG_{[0,\ 1]}\left(s_\tau, n_\tau + \frac{1}{2}r(m-r)\right)$ only by the multiplicative factor $|\beta' P_{\frac{1}{\tau}}\beta|^{-\frac{m}{2}}$.

4 The empirical implementation

To illustrate the above-presented methods the analysis of the price-wage nexus in Poland will be conducted. The data consists of seasonally unadjusted quarterly observations covering the period 1995q1 - 2012q4. The following variables will be analysed:

[w] - the log of nominal wages,

[p] - the log of consumer price index,

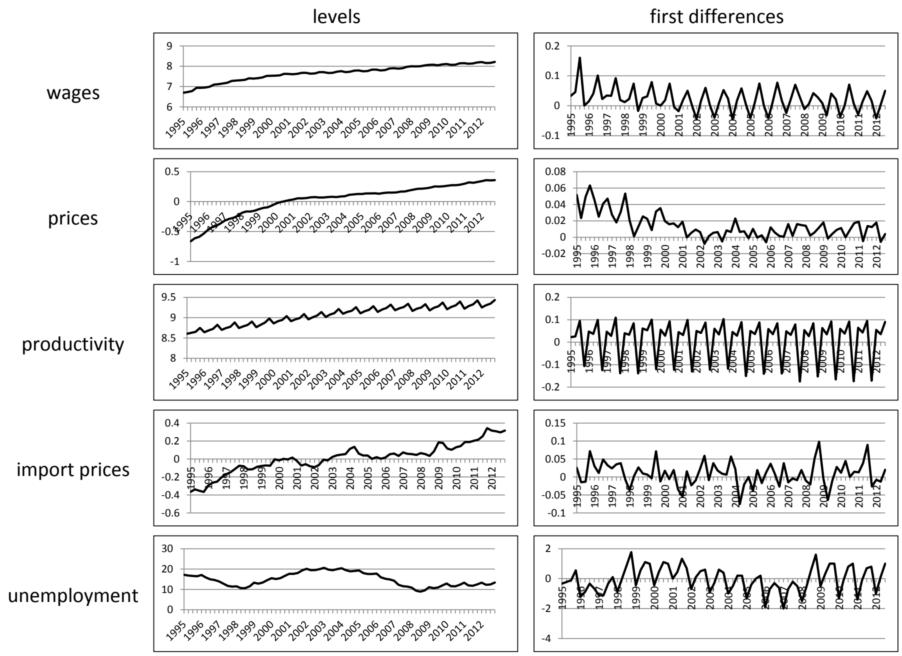
[z] - the log of productivity,

[m] - the log of import prices,

[U] - the unemployment rate.

Lower case letters denote natural logarithms of the original variables. Figure (1) presents the time paths of levels and first differences of the considered variables.
The observed data seems to be the realisations of I(1) processes, appearing to move together, so they may be cointegrated. Previous analyses of a similar set of data within the framework of VEC models with weak form reduced rank structures (VEC-WF) proved that the first differences also display similar behaviour (see Wróblewska 2011, 2012). In this empirical exercise we want to verify the hypothesis of strong

Figure 1: The analysed data



co-movement form. The set of VEC-SF models will be compared with the set of VEC-WF and VEC models.

Additionally, the visual inspection of the analysed variables inclines the researcher to take the possibility of I(2) behaviour into account. In that case the posterior probability of models suggested by economic theory i.e. models with two or more cointegrating relations will be lower than expected. The maximal modulus of unrestricted eigenvalues of the companion matrix will be close to unity and it will stay close to one even in the case of lowering the number of cointegrating vectors.

The set of compared models consists of 60 non-nested specifications. The VAR lag length is assumed to be three. The models may differ in the type of a constant ($d \in \{1, 2\}$, $d = 1$ denotes an unrestricted constant and $d = 2$ - a constant restricted to the cointegration space), the number of cointegrating relations ($r \in \{1, 2, 3, 4\}$), the type of additional restrictions imposed on the short-run model parameters (SF stands for a strong form of an additional restriction, WF - for a weak one), the number of the co-feature vectors ($s \in \{1, 2, 3, 4\}$ in the VEC-WF case, $s \in \{1, \ldots, 5-r\}$ in the VEC-SF model, and $s = 0$ in the VEC model). There are 8 VEC models without additional restrictions, 20 VEC-SF models and 32 VEC-WF models in this set. The VEC-WF

specifications will be analysed by the method similar to that proposed by Wróblewska (2011), but to preserve the coherence of models, the identifying restrictions will be imposed in the same way as proposed in this paper for VEC-SF models.
Equal prior probability of each specification is assumed, i.e. $p(M_{k,d,VEC-.F,s,r}) = \frac{1}{60} \approx 0.0167$, so unequal prior probabilities of different types of co-movement are assumed. The set of VEC-WF models is a priori the most probable, whereas the group of pure VEC models is the least probable.
We specify the following priors of the model parameters (the joint priors are truncated by the stability condition):

VEC-SF's specific parameters:

[-] $\beta|r,d \sim MACG(I_m)|_{\tilde{V}_{r,m}}$, $\quad m = n$ in models with $d = 1$ and $m = n+1$ for $d = 2$,

[-] $G^*|q \sim mN(0, \frac{1}{n}I_q, I_n)$, which leads to $\gamma^*|q \sim MACG(I_n)$; we truncate this prior to $\tilde{V}_{q,n}$,

[-] $D^*|\nu, q \sim mN(0, I_q, \nu I_{n(k-1)+r})$,

VEC's specific parameters:
Following Koop, León-González and Strachan (2010), two equivalent specifications will be considered: $\alpha\beta' = AB'$, $(A \in \mathbb{R}^{nr},\ B \in \mathbb{R}^{mr},\ \beta'\beta = I_r)$. Additionally, we assume that β has positive elements in the first row.

[-] $B|r,d \sim mN(0, \frac{1}{m}I_r, I_m)$, which leads to $\beta|r,d \sim MACG(I_m)$, where $\beta = B(B'B)$; we truncate this prior to $\tilde{V}_{r,m}$,

[-] $A|r,\nu_A, \sim mN(0, \nu_A I_r, I_n)$, $\alpha = A(B'B)^{\frac{1}{2}}$,

[-] $\Gamma|k,\nu \sim mN(0, \nu I_{n(k-1)}, I_n)$,

[-] $\nu_A \sim iG(2, 3)$, $\quad E(\nu_A) = 1,\ Var(\nu_A) = 1$,

VEC-WF's specific parameters:

$$\Delta x_t = \alpha\beta' x_{t-1} + \sum_{i=1}^{k-1} \gamma\delta_i' \Delta x_{t-i} + \Phi D_t + \varepsilon_t = \alpha\beta' x_{t-1} + \gamma\delta' z_{2t} + \Phi D_t + \varepsilon_t,$$

where $z_{2t} = \begin{pmatrix} \Delta x_{t-1}', & \Delta x_{t-2}', & \ldots, & \Delta x_{t-k+1}' \end{pmatrix}'$
and $\delta' = \begin{pmatrix} \delta_1', & \delta_2', & \ldots, & \delta_{k-1}' \end{pmatrix}$,

[-] $B|r,d \sim mN(0, \frac{1}{m}I_r, I_m)$, which leads to $\beta|r,d \sim MACG(I_m)$, where $\beta = B(B'B)^{-\frac{1}{2}}$; we truncate this prior to $\tilde{V}_{r,m}$,

[-] $A|r,\nu_A, \sim mN(0, \nu_A I_r, I_n)$, $\alpha = A(B'B)^{\frac{1}{2}}$,

[-] $G|q \sim mN(0, \frac{1}{n}I_q, I_n)$, which leads to $\gamma|q \sim MACG(I_n)$, where $\gamma = G(G'G)^{-\frac{1}{2}}$; we truncate this prior to $\tilde{V}_{q,n}$,

[-] $D|q,k,\nu \sim mN(0, I_n, \nu I_{n(k-1)})$, $\delta = D(G'G)^{\frac{1}{2}}$,

[-] $\nu_A \sim iG(2,3)$, $\quad E(\nu_A)=1,\; Var(\nu_A)=1$,

Common parameters:

[-] $\Sigma \sim iW(I_n, 1+n+1)$,

[-] $\Gamma_s|h,\Sigma \sim mN(0,\Sigma, hI_{l_s})$,

[-] $h \sim iG(2,3)$, $\quad E(h)=1,\; Var(h)=1$,

[-] $\nu \sim iG(2,3)$, $\quad E(\nu)=1, Var(\nu)=1$.

To compare the above-listed models the Savage-Dickey density ratio (SDDR) will be used (see e.g. Verdinelli and Wasserman 1995). As the model:

$$Z_0 = E, \quad E \sim mN(0,\Sigma) \tag{5}$$

is nested within each considered specification, the Bayes factors for contrasting it with other models will be calculated.
Firstly, we have to point the model parameters, whose values restricting to zero leads to specification (5).
As already mentioned, the entire set of models can be divided into three subsets: VEC-SF, VEC-WF and VEC.
In the VEC-SF group we will test the hypothesis $\delta^* = 0$ and $\Gamma_s = 0$ *versus* $\delta^* \neq 0$ or $\Gamma_s \neq 0$; in the VEC-WF group: $\alpha = 0$ and $\delta = 0$ and $\Gamma_s = 0$ *versus* $\alpha \neq 0$ or $\delta \neq 0$ or $\Gamma_s \neq 0$; in the case of VEC models, $\alpha = 0$ and $\Gamma = 0$ and $\Gamma_s = 0$ *versus* $\alpha \neq 0$ or $\Gamma \neq 0$ or $\Gamma_s \neq 0$. Note that we cannot test whether β or γ^* or γ are zero because we have assumed that their columns are orthonormal.
In each case we need to compute the ratio:

$$\frac{p(\omega_0|X)}{p(\omega_0)},$$

where $p(\omega_0|X)$ denotes the value of the posterior density of ω evaluated at ω_0 and $p(\omega_0)$ is the value of its prior density at the same point, ω states for the tested parameter and ω_0 is the specified value. In the present work, ω_0 equals 0.
Testing whether α, δ^* or δ are zero matrices is equivalent to checking whether, respectively, A, D^* or D are null matrices, as we have assumed that β, γ^* and γ have orthonormal columns and

$$\alpha = A(B'B)^{\frac{1}{2}} \text{with} \det(B'B) \neq 0$$

$$\delta^* = D^*(G^{*'}G^*)^{\frac{1}{2}} \text{with} \det(G^{*'}G^*) \neq 0$$

and

$$\delta = D(G'G)^{\frac{1}{2}} \text{with} \det(G'G) \neq 0$$

This correspondence will be used in the ongoing research. We know the full conditional posterior densities of the tested models parameters. They can be numerically marginalised, i.e. we will estimate $p(\omega_0|X)$ by $\hat{p}(\omega_0|X) = \frac{1}{N}\sum_{i=1}^{N} p(\omega_0|X,\psi_i)$, where $\{\psi_i\}_{i=1}^{N}$ denotes the non-tested parameters sample from the posterior distribution (Verdinelli and Wasserman 1995). If needed, the value $p(\omega_0)$ can be estimated in the same way with the use of a sample from the prior distribution.
Table 1 lists the most probable models in each considered group.

Table 1: The most probable models in each analysed group

rank	k	d	$type$	$s=n-q$	r	$p\left(M_{(k,d,type,s,r)}\|X\right)$
1	3	4	WF	4	1	0.452
2	3	4	WF	3	1	0.435
3	3	4	WF	1	1	0.051
4	3	4	WF	2	1	0.038
5	3	4	WF	2	2	0.023
6	3	4	WF	1	2	$7.9\ 10^{-6}$
7	3	4	VEC	0	1	$6.6\ 10^{-6}$
⋮	⋮	⋮	⋮	⋮	⋮	
23	3	4	SF	1	1	$2.7\ 10^{-10}$

Almost all posterior probability mass is gained by the models with the weak form of common cyclical features. This group strongly dominates the others. VEC-SF models turned out to be the worst. It is worth noting that, as expected, in each group, contrary to economic theory, models with only one cointegrating relation match the data best. Additionally, the absolute value of the maximum free eigenvalue of the companion matrix is close to one, so as previously mentioned, it might be a symptom of I(2) behaviour (see e.g. Juselius (2007), pp. 297-302, for a more detailed discussion), but this is only one possible explanation of these alarming results and it should be formally tested (e.g. through the Bayesian model comparison in the enlarged model group). Without additional tests one could only note that each of the considered models probably lacks describing some important features of the analysed data.
As the main aim of this section was to examine how many common cyclical features drive the modelled data, we sum up the obtained results in table (2). One can easily notice that, in the case of additional reduced rank restrictions, models with three or four common cycles gathered almost all the posterior probability mass while in the strong form group models with one co-feature beat the others. As shown by Hecq, Palm and Urbain (2006), the existence of s weak form common feature vectors with s greater than the number of cointegrating vectors (r), implies the existence of $s-r$ strong form common features (see Lemma 2 in Hecq, Palm and Urbain 2006, p. 122). In our case the best model is the one with four weak form co-features and

Table 2: Posterior probabilities of the number of co-feature vectors (marginal and in the SF, WF groups)

$s = n - q$	*type* SF	WF	$p(s\|X)$
4	1.8 10^{-11}	0.452	0.452
3	3.3 10^{-12}	0.435	0.435
2	1.0 10^{-11}	0.061	0.061
1	2.7 10^{-10}	0.051	0.051
$p(type\|X)$	3.038 10^{-10}	≈ 1	≈ 1

one cointegrating relation, so it suggests the presence of three strong form common feature vectors, more than is implied by the most probable VEC-SF model, so it may be worth expanding the compared group of models with mixed form common features. As it goes beyond the main aim of this paper, it is left for future research.

5 Concluding remarks

Based on the idea of common serial correlation feature introduced by Engle nad Kozicki (1993) and combined with the notion of cointegration by Vahid and Engle (1993) (see also Hecq, Palm and Urbain 2006), this paper proposed a Bayesian counterpart for the analysis of this concept. It should be emphasised that owing to the idea of Bayesian model comparison one is free of deciding whether one should at first test the number of cointegrating relations or one should start with the tests for the number and type of short-run restrictions (see the discussions in Hecq, Palm and Urbain 2006, and Athanasopoulos, Guillén de Carvalho, Issler and Vahid 2011 among others).

The proposed methods were illustrated by the analysis of the price-wage nexus in the Polish economy based on the quarterly data spanning the period 1995-2012. The data supported co-feature restrictions, but of the weak rather than the strong form.

Finally, it is worth noting that imposing additional lower rank restrictions could affect further results which could be obtained in the framework of the VEC model, such as forecasting (see e.g. Vahid and Issler 2002, and Athanasopoulos, Guillén de Carvalho, Issler and Vahid 2011), impulse response analysis, permanent-transitory decomposition. In our example the posterior probability of the VEC-SF model was negligible so we decided to postpone such an analysis for other researches.

6 Acknowledgements

I would like to thank the anonymous Referee for careful assessment of my work, valuable comments and suggestions which helped to correct some significant errors

that appeared in preliminary versions of this paper.
I am also thankful to Jacek Osiewalski and Łukasz Kwiatkowski for the time devoted to me, invaluable discussions and comments.
All errors and omissions are mine.
Publication was financed from the funds granted to the Faculty of Management (Department of Econometrics and Operational Research) at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

[1] Abbas A.E., Bakir N.O., Klutke G-A., Sun Z. (2013), Effects of risk aversion on the value of information in two-action decision problems, *Decision Analysis* 10(3), 257–275.

[2] Athanasopoulos G., Guillén O.T. de Carvalho, Issler J.V., Vahid F. (2011), Model selection, estimation and forecasting in VAR models with short-run and long-run restrictions, *Journal of Econometrics* 164, 116–129.

[3] Byrne S. Girolami M. (2013), Geodesic Monte Carlo on embedded manifolds, *Scandinavian Journal of Statistics* 40, 825–845.

[4] Chikuse Y. (2002), *Statistics on special manifolds*, *Lecture Notes in Statistics*, vol. 174, Springer-Verlag, New York.

[5] De Wall D.J. (1979), On the normalizing constant for the Bingham-von Mises-Fisher matrix distribution, *South African Statistical Journal* 13, 103–112.

[6] Engle R.F., Kozicki S. (1993), Testing for common features, *Journal of Business and Economic Statistics* 11, 369–380.

[7] Ericsson N.R. (1993), Comment (to the paper Testing for common features by Engle and Kozicki), *Journal of Business and Economic Statistics* 11, 380–383.

[8] Hecq A., Palm F.C., Urbain J.P. (2006), Common cyclical features analysis in VAR models with cointegration, *Journal of Econometrics* 132, 117–141.

[9] Hoff P.D. (2009), Simulation of the Matrix Bingham-von Mises-Fisher Distribution, with applications to multivariate and relation data, *Journal of Computational and Graphical Statistics* 18, 438–456.

[10] James (1954), Normal multivariate analysis and the orthogonal group, *Annals of Mathematical Statistics* 25, 40–75.

[11] Juselius K. (2007), *The Cointegrated VAR Model. Methodology and Applications*, Oxford University Press, second edition

[12] Khatri C.G., Mardia K.V. (1977), The von Mises-Fisher matrix distribution in orientation statistics, *Journal of the Royal Statistical Society*, Series B, 39, 95–106.

[13] Koop G., León-González R., Strachan R. (2010), Efficient posterior simulation for cointegrated models with priors on the cointegration space, *Econometric Reviews* 29, 224–242.

[14] Strachan R., Inder (2004), Bayesian analysis of the error correction model, *Journal of Econometrics* 123, 307-0325.

[15] Villani M. (2005), Bayesian reference analysis of cointegration, *Econometric Theory* 21, 326–357.

[16] Vahid F., Engle R.F. (1993), Common trends and common cycles, *Journal of Applied Econometrics* 8, 341–360.

[17] Vahid F., Issler J.V. (2002), The importance of common cyclical features in VAR analysis: a Monte-Carlo study, *Journal of Econometrics* 109, 341–363.

[18] Verdinell I., Wasserman L. (1995), Computing Bayes factors using a generalization of the Savage-Dickey Density Ratio, *Journal of the American Statistical Association* 90, 614–618.

[19] Wróblewska J. (2011), Bayesian analysis of weak form reduced rank structure in VEC models, *Central European Journal of Economic Modelling and Econometrics* 3, 169–186.

[20] Wróblewska J. (2012), Bayesian analysis of weak form polynomial reduced rank structures in VEC models, *Central European Journal of Economic Modelling and Econometrics* 4, 253–267.

[21] Zellner A. (1971), *An introduction to Bayesian inference in econometrics*, J. Wiley, New York.

Appendix A – The full conditional posteriors of model parameters

The full conditional posteriors are proportional to the joint posterior density function, which is obtained as a product of the joint prior and the likelihood function. It has the following kernel:

$$p(\Sigma, \Gamma_s, G^*, D^*, \beta, \nu, h, \tau, \tau^*|X) \propto$$

$$\Sigma|^{-\frac{1}{2}(q_\Sigma+l_s+T+n+1)}|P_\tau|^{-\frac{r}{2}}|\beta' P_\tau^{-1}\beta|^{-\frac{m}{2}}|P_{\tau^*}|^{-\frac{q}{2}}\times$$

$$\times h^{-\frac{nl_s}{2}-n_h-1}\nu^{-\frac{[n(k-1)+r]q}{2}-n_\nu-1}\exp\left(-\frac{s_h}{h}\right)\exp\left(-\frac{s_\nu}{\nu}\right)\times$$

$$\times\exp\left\{-\frac{1}{2}tr\left[\Sigma^{-1}\left(S+\frac{1}{h}\left(\Gamma_s-\underline{\mu}_s\right)'\left(\Gamma_s-\underline{\mu}_s\right)E'E\right)\right]\right\}\times$$

$$\times\exp\left\{-\frac{1}{2}tr\left(nG^{*'}P_{\tau^*}^{-1}G^*\right)\right\}\exp\left\{-\frac{1}{2}tr\left[\frac{1}{\nu}\underline{\Omega}_{D^*}^{-1}\left(D^*-\underline{\mu}_{D^*}\right)'\left(D^*-\underline{\mu}_{D^*}\right)\right]\right\}\times$$

$$\times\tau^{*-n_{\tau^*}-1}exp\left(-\tfrac{s_{\tau^*}}{\tau^*}\right)\mathbf{1}_{[0,1]}(\tau^*)\tau^{-n_\tau-1}exp\left(-\tfrac{s_\tau}{\tau}\right)\mathbf{1}_{[0,1]}(\tau),$$

where $E = Z_0 - \begin{pmatrix} Z_1\beta, & Z_2\end{pmatrix} D^*G^{*'} - Z_3\Gamma_s$, $\mathbf{1}_{[0,1]}(a)$ denotes a function which takes 1 for a in $[0,\ 1]$ and 0 elsewhere, l_s stands for the column dimension of Γ_s and $[d\beta]$ denotes the normalised invariant measure on the Stiefel manifold (see e.g. James 1954).
It is straightforward to see that the full conditional posterior of Σ is the inverted Wishart distribution, i.e.

$$p(\Sigma|., X) = f_{iW}\left(S+\frac{1}{h}\left(\Gamma_s-\underline{\mu}_s\right)'\left(\Gamma_s-\underline{\mu}_s\right)+E'E, q_\Sigma+l_s+T\right).$$

It is also obvious that the full conditional posteriors of ν and h are inverted Gamma distributions:

$$p(\nu|., X) = f_{iG}\left(s_\nu+\frac{1}{2}tr\left[\underline{\Omega}_{D^*}^{-1}\left(D^*-\underline{\mu}_{D^*}\right)'\left(D^*-\underline{\mu}_{D^*}\right)\right], n_\nu+\frac{[n(k-1)+r]q}{2})\right)$$

and

$$p(h|., X) = f_{iG}\left(s_h+\frac{1}{2}tr\left[\Sigma^{-1}\left(\Gamma_s-\underline{\mu}_s\right)'\left(\Gamma_s-\underline{\mu}_s\right)\right], n_h+\frac{nl_s}{2}\right)$$

The full conditional distribution of τ is proportional to

$$|P_\tau|^{-\frac{r}{2}}|\beta' P_{\frac{1}{\tau}}\beta|^{-\frac{m}{2}}\exp\left(-\frac{s_\tau}{\tau}\right)\tau^{-n_\tau-1}\mathbf{1}_{[0,1]}(\tau).$$

Let us recall that $P_\tau = H_B H_B' + \tau H_{B\perp} H_{B\perp}'$ with $H_B' H_B = I_r$, $H_{B\perp}' H_{B\perp} = I_{m-r}$ and $H_B' H_{B\perp} = \mathbf{0}$, so $H_B H_B'$ is a projection matrix on the space spanned by H_B ($sp(H_B)$) and $H_{B\perp} HB \perp'$ is a projection matrix on the orthogonal complement of $sp(H_B)$, then $H_{B\perp} H_{B\perp}' = I_m - H_B H_B'$. Using the last relation we get

$$\begin{aligned} |P_\tau| &= |H_B H_B' + \tau(I_m - H_B H_B')| = |H_B H_B' + \tau I_m - \tau H_B H_B'| \\ &= |(1-\tau) H_B H_B' + \tau I_m| = \tau^m |\frac{1-\tau}{\tau} H_B H_B' + I_m| \\ &= \tau^m |\frac{1-\tau}{\tau} H_B' H_B + I_r| = \tau^m |\frac{1-\tau}{\tau} I_r + I_r| = \tau^m |\frac{1}{\tau} I_r| \\ &= \tau^{m-r}. \end{aligned}$$

Finally, the full conditional posterior of τ is proportional to

$$|\beta' P_{\frac{1}{\tau}} \beta|^{-\frac{m}{2}} \exp\left(-\frac{s_\tau}{\tau}\right) \tau^{-n_\tau - \frac{1}{2} r(m-r) - 1} \mathbf{1}_{[0,1]}(\tau),$$

which is non-standard.
Using the very same calculations for $|P_{\tau^*}|$ as for $|P_\tau|$, we obtain the full conditional posterior distribution of τ^*:

$$p(\tau^*|., X) = iG_{[0,1]} \left(s_{\tau^*} + \frac{1}{2} tr\left(D^{*'} H_{D^*\perp} H_{D^*\perp}' D^* \right), n_{\tau^*} + \frac{1}{2} q[n(k-1) + r - q] \right).$$

Adopting the commonly known calculations for multivariate regressions (see e.g. Zellner 1971, pp. 224-227), we get the full conditional posterior of Γ_s as

$$mN\left(\overline{\mu}_s, \Sigma, (\frac{1}{h} I_{l_s} + Z_3' Z_3)^{-1} \right),$$

where

$$\overline{\mu}_s = \left(\frac{1}{h} I_{l_s} + Z_3' Z_3 \right)^{-1} \left[\frac{1}{h} \underline{\mu}_s + Z_3' \left(Z_0 - \begin{pmatrix} Z_1\beta, & Z_2 \end{pmatrix} D^* G^{*'} \right) \right].$$

The full conditional posteriors of D^* (δ^*) and G^* (γ^*) could be obtained with the methods presented and discussed in Koop, León-González and Strachan (2010) in the context of the VEC model.
To get the full conditional posterior of β let us denote the last $n(k-1)$ rows of the matrix $D^* G^{*'}$ by $\overline{D^* G^{*'}}$ and the first r rows of $D^* G^{*'}$ by $\tilde{\Gamma}$. Rewrite $\exp\{-\frac{1}{2} tr[\Sigma^{-1} E' E]\}$ as $\exp\{-\frac{1}{2} tr[\Sigma^{-1} (\tilde{Z} - Z_1 \beta \tilde{\Gamma})' (\tilde{Z} - Z_1 \beta \tilde{\Gamma})]\}$, where

$$\tilde{Z} = Z_0 - Z_2 \overline{D^* G^{*'}} - Z_3 \Gamma_s$$

The last expression is proportional to

$$\exp\left\{-\frac{1}{2}tr\left(-2\Sigma^{-1}\tilde{Z}'Z_1\beta\tilde{\Gamma}+\Sigma^{-1}\tilde{\Gamma}'\beta'Z_1'Z_1\beta\tilde{\Gamma}\right)\right\}=$$
$$=\exp\left\{tr\left(\tilde{\Gamma}\Sigma^{-1}\tilde{Z}'Z_1\beta-\frac{1}{2}\tilde{\Gamma}\Sigma^{-1}\tilde{\Gamma}'\beta'Z_1'Z_1\beta\right)\right\}.$$

Finally,

$$p\left(\beta|\Sigma,G^*,D^*\Gamma_s,q,h,\nu\right)\propto\exp\left\{tr\left(F'\beta+\tilde{B}\beta'\tilde{A}\beta\right)\right\}[d\beta],$$

where $F=Z_1'\tilde{Z}\Sigma^{-1}\tilde{\Gamma}'$, $\tilde{A}=Z_1'Z_1$, $\tilde{B}=-\frac{1}{2}\tilde{\Gamma}\Sigma^{-1}\tilde{\Gamma}'$. The obtained function is the kernel of the Bingham-von Mises-Fisher distribution.