Econometrics and Statistics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Bayesian analysis of seasonally cointegrated VAR models

Justyna Wróblewska

Department of Econometrics and Operational Research, Krakow University of Economics, Rakowicka 27, Kraków, 31–510, Poland

ARTICLE INFO

Article history:
Received 27 November 2021
Revised 11 February 2023
Accepted 11 February 2023
Available online xxx

JEL classification:
C11
C32
C53

Keywords:
Seasonal cointegration
Reduced rank regression
Error correction model
Bayesian model comparison

ABSTRACT

The aim is to develop a Bayesian seasonally cointegrated model for quarterly data. Relevant prior structure is proposed, and the set of full conditional posterior distributions is derived, enabling us to employ the Gibbs sampler for posterior inference. The identification of cointegrating spaces is obtained by orthonormality restrictions imposed on vectors spanning them. The point estimation of the cointegrating spaces is also discussed. In the presence of a seasonal pattern with one cycle per year, the cointegrating vectors belong to the complex space, which should be taken into account in the identification scheme. The methodology is illustrated by the analysis of money and prices in the Polish economy.

1. Introduction

Many macroeconomic time series, quarterly or monthly, display both strong trend and seasonal pattern. The idea of cointegration at zero frequency (assuming a common stochastic trend controlling the long-run behavior of the series) is well-known and often employed in empirical analyses. Methods enabling the estimation of parameters of vector error correction models are already well-developed within both the frequentist (e.g., Johansen, 1995) and Bayesian framework (e.g., Koop et al., 2006).

However, although the idea of cointegration at seasonal frequencies was introduced already in 1990 (Hylleberg et al., 1990), it has not been very popular in the literature. In most of applied studies, either the seasonally adjusted data are analyzed, or the seasonality is accounted for by a set of seasonal dummies. However, some papers present quite undesirable results of seasonal adjustments. Such procedures may for example change both the short- as well as long-run behavior of the series (e.g., Nerlove, 1964; Cubadda, 1999; Hecq, 1998; Granger and Siklos, 1995), similarly as in the case of filtering nonseasonal unit roots (e.g., Meyer and Winker, 2005; Hamilton, 2018).

Abeysinghe (1994) points to some problems which may occur if stochastic seasonal behavior is modeled by deterministic dummies. The major problem is the risk of spurious correlations. It is also shown that seasonally integrated processes can generate a regular seasonal pattern, so even series with deterministic periodic behavior can be approximated by processes with stochastic seasonality. Abeysinghe (1994) also finds evidence that it is better to employ seasonal filtering rather than resort to dummy variables, although the filtered series typically requires higher AR lags. Kunst (1993) analyzes data from four European economies and shows that the sources of seasonal fluctuations vary over the countries under study.

E-mail address: wroblewj@uek.krakow.pl

https://doi.org/10.1016/j.ecosta.2023.02.002
2452-3062/

In view of the above-mentioned risks stemming from a poor choice of the approach to handling seasonality, it seems of utmost importance to carefully investigate its sources, which in more formal terms translates into the problem of model comparison. Among its other merits, the Bayesian statistical framework appears particularly conducive in that regard, as it allows not only for a formal and fully probabilistic comparison of models featuring different seasonal structures, but also to formally combine posterior inference from these various models within the pooling approach. It seems even more valid an argument, as the papers quoted above show that it is often difficult to pinpoint a single source of seasonal variability. As mentioned above, the researcher is not 'stuck' with the seasonal structure pre-specified in a given sampling model, but can actually merge posterior results stemming from models featuring different natures of seasonal fluctuations, to form 'final' posterior distributions pertaining to, for example, structural analyses or forecasting.

The predictive performance of VAR models for seasonal time series is discussed by, among others, Reimers (1997), Kunst and Franses (1998), Löf and Franses (2001). Reimers (1997), in an extensive simulation study, shows that models for the first differences, featuring seasonal dummies, outperform models with seasonal cointegration in forecasting for shorter horizons (up to four quarters ahead), whereas for longer horizons, models for the fourth differences tend to produce more accurate point forecasts. Kunst and Franses (1998) deliver similar findings through a forecasting exercise performed for real income, consumption and wealth in the German economy. Kunst and Franses (1998) analyze three data sets and three sets of models: without seasonal dummies, with dummies restricted to the cointegrating vectors, and with unrestricted dummies. The paper discusses in detail the role that seasonal dummies play for inference about the number of cointegrating relations and (as a result) also for forecasting. The results differ across various data sets and forecast horizons, but it still can be concluded that, in general, models with unrestricted seasonal constants are usually outperformed by the alternative specifications. Löf and Franses (2001) show that handling seasonal patterns in data through seasonal differences results in lower root mean square forecast errors for horizons longer than one period ahead. Moreover, among the VEC specifications considered in the cited work (seasonal VECs as well as nonseasonal VECs featuring seasonally varying intercepts), it is the seasonally cointegrated models that deliver superior forecasts. In general, these results appear to validate making strides in development of methodology for the seasonal VEC models.

So far, the seasonal VEC specifications have been considered solely within the frequentist statistical framework. The likelihood methods enabling the estimation of seasonal VECs were proposed by Johansen and Schaumburg (1999) (see also Cubadda and Omtzigt (2005) for further discussion). However, to the best of our knowledge, no Bayesian estimation and inference framework has been developed for the models in question. This paper aims to fill this gap by formulation of a Bayesian VEC model for quarterly data with seasonal cointegration. To that end, a relevant prior distribution for the parameters is proposed, taking into account the geometry of VEC models. Then, we derive the set of full conditional posteriors that underlie the Gibbs sampler for posterior inference. Additionally, the point estimation of cointegrating spaces is also discussed. The methodology is illustrated by an empirical analysis of money and prices in the Polish economy, modeled within a four-dimensional seasonal system. We begin the analysis with Bayesian model comparison to identify the model featuring the most adequate seasonal structure. Then, we proceed with a comparison of the forecasting performance within the winning specifications.

2. Bayesian seasonal VEC model

The original definition of cointegration was formulated by Granger (1981) for VAR processes integrated at zero frequency, i.e. driven by stochastic trends and having unit roots at zero frequency. Non-stationary series are said to be cointegrated when there exist non-zero linear combinations of the series which lower the order of their integration. In such a case, the series share the stochastic trends, so they are driven by the same persistent shocks and move together in the long run (Juselius, 2006). In their seminal paper, Engle and Granger (1987) analyzed further the relation between cointegration and error correction. In the case of vector error correction models, the presence of stochastic trends leads to a reduced-rank restriction imposed on the long-run matrix (see, e.g., Johansen (1995), and Juselius (2006) for further details). However, also seasonal behavior of economic time series may be driven by (seasonally) integrated processes, featuring unit roots at seasonal frequencies (e.g., Cubadda (1999), and the references therein). If these processes are cointegrated at these frequencies, then they share common stochastic seasonal components (Hylleberg et al., 1990), so there exist non-zero linear combinations which cancel some seasonal unit roots. Hylleberg et al. (1990) discussed error correction representations of series cointegrated at various frequencies. Similarly as in the zero frequency case, the presence of seasonal cointegration leads to reduced-rank restrictions imposed on the VEC parameters connected with a given seasonal frequency (see also Johansen and Schaumburg, 1999).

To specify a quarterly, seasonally cointegrated VAR(k) model for an n-dimensional series $\{y_t\}$, let us write its basic form first:

$$y_t = \sum_{j=1}^{k} A_j y_{t-j} + \Phi D_t + \varepsilon_t, \quad \varepsilon_t \sim iidN(0, \Sigma), \tag{1}$$

where D_t contains deterministic components such as a constant, trend or seasonal dummies. In particular, we consider a deterministic component of the form: $\Phi D_t = \mu + a\cos\left(\frac{\pi}{2}t\right) + b\sin\left(\frac{\pi}{2}t\right) + c\cos(\pi t) + \gamma t$ (Franses and Kunst, 1999; Kotłowski, 2005). The initial conditions $(y_0, y_{-1}, \ldots, y_{-k+1})$ are fixed and set as pre-sample observations.

According to the Lagrange expansion, the polynomial matrix $A(z) = I_n - A_1 z - \cdots - A_k z^k$ of process (1) may be developed around points $z_1, z_2, \ldots, z_S$ (for a given number S of unit roots) as follows (Johansen and Schaumburg, 1999; Kotłowski, 2005):

$$A(z) = p(z)I_n + \sum_{s=1}^{S} A(z_s)\frac{p_s(z)z}{p_s(z_s)z_s} + p(z)zA_0(z), \tag{2}$$

where $p(z) = \prod_{s=1}^{S}(1 - \bar{z}_s z)$, $p_j(z) = \prod_{s\neq j}^{S}(1 - \bar{z}_s z) = \frac{p(z)}{1-\bar{z}_j z}$, $z \neq z_j$, $A_0(z)$ is a polynomial matrix, and the bar on the top of complex quantities (e.g. $\bar{z}$) denotes the complex conjugate. If z_s is additionally a root of the characteristic polynomial of process (1), i.e. $|A(z_s)| = 0$, than $A(z_s)$ is of reduced rank and can be decomposed as the product of full column rank matrices $A(z_s) = a_s b_s'$.

A general representation of the complex error correction model is discussed, e.g., by Johansen and Schaumburg (1999), Cubadda (2001), Ahn et al. (2004). In this paper, we focus solely on such processes that are cointegrated at zero and quarterly frequencies (i.e. $|A(z)| = 0$ for $z_1 = 1$, $z_2 = -1$, $z_3 = i$, $z_4 = \bar{z}_3 = -i$). The expansion in Eq. (2) leads to the following seasonal error correction representation allowing for a different number of cointegrating relations at each frequency (e.g., Johansen and Schaumburg, 1999; Cubadda and Omtzigt, 2005; Kotłowski, 2005):

$$\Delta_4 y_t = \alpha_1\beta_1'\tilde{y}_t^{(1)} + \alpha_2\beta_2'\tilde{y}_t^{(2)} + \alpha_\star\bar{\beta}_\star'\tilde{y}_t^{(3)} + \bar{\alpha}_\star\beta_\star'\bar{\tilde{y}}_t^{(3)} + \sum_{i=1}^{k-4}\Gamma_j\Delta_4 y_{t-j} + \tilde{\Phi}\tilde{D}_t + \varepsilon_t, \quad \varepsilon_t \sim iiN(0,\Sigma), \tag{3}$$

where $\Delta_4 y_t = (1 - L^4)y_t = y_t - y_{t-4}$ with L denoting the lag operator, and $\Gamma_j = -\sum_{l=1}^{[(k-j)/4]} A_{j+4l}$, $j = 1, 2, \ldots, k-4$. Vectors $\tilde{y}_t^{(\cdot)}$ may contain also deterministic components, so that $\tilde{y}_t^{(\cdot)} = \begin{pmatrix} y_t^{(\cdot)} \\ d_t^{(\cdot)} \end{pmatrix}$.

Applying polynomials $p_j(L)$ to stochastic and deterministic parts of process (1) results in vectors $\tilde{y}_t^{(\cdot)}$ of the following form:

- at zero frequency (long-run behavior, zero cycles per year): $\tilde{y}_t^{(1)} = \begin{pmatrix} y_t^{(1)} \\ d_t^{(1)} \end{pmatrix}$, where $y_t^{(1)} = p_1(L)Ly_t = (1 + L + L^2 + L^3)Ly_t = y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4}$, while the term $p_1(L)LD_t = (1 + L + L^2 + L^3)LD_t$ leads to $d_t^{(1)} = t - \frac{5}{2}$ and an unrestricted constant,
- at π (bi-annual) frequency (two cycles per year): $\tilde{y}_t^{(2)} = \begin{pmatrix} y_t^{(2)} \\ d_t^{(2)} \end{pmatrix}$, where $y_t^{(2)} = p_2(L)Ly_t = (1 - L + L^2 - L^3)Ly_t = y_{t-1} - y_{t-2} + y_{t-3} - y_{t-4}$, while the term $p_2(L)LD_t = (1 - L + L^2 - L^3)LD_t$ leads to $d_t^{(2)} = -cos(\pi t)$ and an unrestricted constant,
- at $\frac{\pi}{2}$ (annual) frequency (one cycle per year): $\tilde{y}_t^{(3)} = \begin{pmatrix} y_t^{(3)} \\ d_t^{(3)} \end{pmatrix}$, where $y_t^{(3)} = p_3(L)Ly_t = (-i - L + iL^2 + L^3)Ly_t = -iy_{t-1} - y_{t-2} + iy_{t-3} + y_{t-4}$, while the term $p_3(L)LD_t = (-i - L + iL^2 + L^3)LD_t$ leads to $d_t^{(3)} = \begin{pmatrix} \cos\left(\frac{\pi}{2}t\right) - i\sin\left(\frac{\pi}{2}t\right) \\ -\sin\left(\frac{\pi}{2}t\right) - i\cos\left(\frac{\pi}{2}t\right) \end{pmatrix}$ and an unrestricted constant.

Note that the unrestricted constants, occurring for each of the above frequencies, result from the linear trend assumed for the level of the analyzed process. If there is no linear trend, and the process features only a constant, there is neither a trend restricted to the cointegration space at zero frequency nor an unrestricted constant in the representation given by (3), but there is a constant restricted to the cointegration space at zero frequency ($d_t^{(1)} = 1$); see Juselius (2006, pp. 93-112) for an in-depth discussion of the meaning of dummies gathered in vectors D_t and $\tilde{D}_t$, and also the relations between them in the VAR model cointegrated at zero frequency.

Returning to Eq. (3), note that $\alpha_\star\bar{\beta}_\star'\tilde{y}_t^{(3)}$ and $\bar{\alpha}_\star\beta_\star'\bar{\tilde{y}}_t^{(3)}$ are complex conjugate matrices, so they sum to a double of their real part:

$$\alpha_\star\bar{\beta}_\star'\tilde{y}_t^{(3)} + \bar{\alpha}_\star\beta_\star'\bar{\tilde{y}}_t^{(3)} = 2Re(\alpha_\star\bar{\beta}_\star'\tilde{y}_t^{(3)}) = 2\Big[(\alpha_I\beta_R' - \alpha_R\beta_I')(y_{t-1} - y_{t-3}) - (\alpha_R\beta_R' + \alpha_I\beta_I')(y_{t-2} - y_{t-4})\Big], \tag{4}$$

where α_R and β_R denote the real parts of $\alpha_\star$ and $\beta_\star$, respectively, whereas α_I and β_I - their imaginary parts ($\alpha_\star = \alpha_R + i\alpha_I$, $\beta_\star = \beta_R + i\beta_I$).

Transformations presented in (4) lead to a more common representation of seasonally cointegrated quarterly VAR process (e.g., Johansen and Schaumburg, 1999; Cubadda and Omtzigt, 2005; Kotłowski, 2005):

$$\Delta_4 y_t = \Pi_1\tilde{y}_t^{(1)} + \Pi_2\tilde{y}_t^{(2)} + \Pi_3\tilde{y}_t^{(32)} + \Pi_4\tilde{y}_t^{(31)} + \sum_{i=1}^{k-4}\Gamma_i\Delta_4 y_{t-i} + \tilde{\Phi}\tilde{D}_t + \varepsilon_t, \quad \varepsilon_t \sim iiN(0,\Sigma), \tag{5}$$

where:

- $\Pi_1 = \alpha_1\beta_1'$, $\Pi_2 = \alpha_2\beta_2'$, $\Pi_3 = -2(\alpha_R\beta_R' + \alpha_I\beta_I')$, $\Pi_4 = 2(\alpha_I\beta_R' - \alpha_R\beta_I')$,

- $\tilde{y}_t^{(31)} = \begin{pmatrix} y_t^{(31)} \\ d_t^{(31)} \end{pmatrix}$, $y_t^{(31)} = (1-L^2)Ly_t = y_{t-1} - y_{t-3}$, $d_t^{(31)} = \begin{pmatrix} -\cos(\frac{\pi t}{2}) \\ \sin(\frac{\pi t}{2}) \end{pmatrix}$,
- $\tilde{y}_t^{(32)} = \begin{pmatrix} y_t^{(32)} \\ d_t^{(32)} \end{pmatrix}$, $y_t^{(32)} = (1-L^2)L^2y_t = y_{t-2} - y_{t43}$, $d_t^{(32)} = \begin{pmatrix} -\sin(\frac{\pi t}{2}) \\ -\cos(\frac{\pi t}{2}) \end{pmatrix}$.

Since modeled data carry information only about the cointegration space and not the cointegration vectors, in estimation of the model one needs to deal with the non-identification occurring in the products: $\alpha_1\beta_1'$, $\alpha_2\beta_2'$, $\bar{\beta}_\star\alpha_\star'$, $\beta_\star\bar{\alpha}_\star'$. To that end, we employ methods proposed by Koop et al. (2009) and adapt them to the complex case. Following their ideas, we consider two observationally equivalent representations for each of the products mentioned above. In their $A-B$ representations it is assumed that matrices A and B belong to the $\mathbb{R}^{\cdot}$ or $\mathbb{C}^{\cdot}$ spaces of appropriate dimensions, whereas in the $\alpha-\beta$ representations, βs have orthonormal columns, while αs belong to the $\mathbb{R}^{\cdot}$ or $\mathbb{C}^{\cdot}$ spaces:

- $A_1B_1' \equiv \alpha_1\beta_1'$,
  $A_1 \in \mathbb{R}^{n\times r_1}, B_1 \in \mathbb{R}^{m_1\times r_1}, \alpha_1 = A_1(B_1'B_1)^{\frac{1}{2}} \in \mathbb{R}^{n\times r_1}, \beta_1 = B_1(B_1'B_1)^{-\frac{1}{2}} \in \mathbb{V}_{r_1,m_1}$,
- $A_2B_2' \equiv \alpha_2\beta_2'$,
  $A_2 \in \mathbb{R}^{n\times r_2}, B_2 \in \mathbb{R}^{m_2\times r_2}, \alpha_2 = A_2(B_2'B_2)^{\frac{1}{2}} \in \mathbb{R}^{n\times r_2}, \beta_2 = B_2(B_2'B_2)^{-\frac{1}{2}} \in \mathbb{V}_{r_2,m_2}$,
- $A_\star\bar{B}_\star' \equiv \alpha_\star\bar{\beta}_\star'$,
  $A_\star = A_R + iA_I \in \mathbb{C}^{n\times r_3}, B_\star = B_R + iB_I \in \mathbb{C}^{m_3\times r_3}, \alpha_\star = A_\star(\bar{B}_\star'B_\star)^{\frac{1}{2}} \in \mathbb{C}^{n\times r_3}, \beta_\star = B_\star(\bar{B}_\star'B_\star)^{-\frac{1}{2}} \in \mathbb{V}^{\mathbb{C}}{}_{r_3,m_3}$,

where $\mathbb{V}_{r_j,m_j}$, $j=1,2$ denotes the Stiefel manifold, i.e. the set of $m_j\times r_j$ matrices with orthonormal columns ($\mathbb{V}_{r_j,m_j} = \left\{X \in \mathbb{R}^{m_j\times r_j} : X'X = I_{r_j}\right\}$), and $\mathbb{V}^{\mathbb{C}}{}_{r_3,m_3}$ stands for the complex Stiefel manifold, i.e. the set of $m_3\times r_3$ semi-unitary matrices ($\mathbb{V}^{\mathbb{C}}{}_{r_3,m_3} = \left\{X \in \mathbb{C}^{m_3\times r_3} : \bar{X}'X = I_{r_3}\right\}$).

Note that the above approach solves the non-identification issue only partially as there is a many-to-one relationship between the Stiefel manifolds and the Grassmann manifolds to the latter of which belong cointegration spaces. The Grassmann manifolds ($\mathbb{G}_{r_j,m_j-r_j}, j=1,2$, $\mathbb{G}^{\mathbb{C}}{}_{r_3,m_3-r_3}$) collect r_j-dimensional planes passing through the origin, in the real ($j=1,2$) or complex ($j=3$) vector m_j-dimensional space (e.g., James, 1954; Chern and Wolfson, 1987). If X is an element of the (complex) Stiefel manifold, and a $r_j\times r_j$ ($j=1,2,3$) matrix O is an element of the (complex) orthonormal group ($O'O = OO' = I_{r_j}$, $j=1,2$, $\bar{O}'O = O\bar{O}' = I_{r_3}$), then XO also belongs to the (complex) Stiefel manifold, and both, X and XO, span the same spaces (the projection matrices are equal, i.e. $XX' = XOO'X'$ in the real case, and $X\bar{X}' = XO\bar{O}'\bar{X}'$ in the complex case).

To handle the non-identification related to the first two of the above-mentioned products (i.e. $\alpha_1\beta_1'$ and $\alpha_2\beta_2'$), involving only real matrices, one can directly apply the approach developed by Koop et al. (2009). The third product requires a straightforward adaptation of the discussed methods for the case of complex matrices and spaces (details are discussed below the following presentation of the prior distributions and also in the Appendix).

Formulation of a Bayesian statistical model requires specification of the prior distribution of the model's parameters. We impose the following prior structure in the $A-B$ parameterization:

- the inverse Wishart distribution for the covariance matrix: $\Sigma \sim iW(S,q)$,
- if the researcher is willing to estimate S (the matrix hyperparameter in prior distribution for Σ) Wishart prior can be employed: $S \sim W(A_S, q_s)$,
- the matrix normal distribution for Γ: $\Gamma|\Sigma,\nu \sim MN(\underline{\mu}_\Gamma, \Sigma, \nu\underline{\Omega}_\Gamma)$,
- the matrix normal distribution for the adjustment coefficients at frequency 0: $A_1|\Sigma,\nu \sim MN(\underline{\mu}_1, \nu\underline{\Omega}_1, \Sigma)$,
- the matrix normal distribution for the adjustment coefficients at frequency π: $A_2|\Sigma,\nu \sim MN(\underline{\mu}_2, \nu\underline{\Omega}_2, \Sigma)$,
- the complex matrix normal distribution for the adjustment coefficients at frequency $\frac{\pi}{2}$: $A_\star|\Sigma,\nu \sim mCN(\underline{\mu}_\star, \nu I_{r_3}, \Sigma)$, i.e. $p(A_\star|\Sigma) = \pi^{-nr_3}|\Sigma|^{-r_3}\exp\{-tr\Sigma^{-1}(A_\star - \underline{\mu}_\star)(\frac{1}{\nu}I_{r_3})(\bar{A}_\star - \bar{\underline{\mu}}_\star)'\}$ (e.g., Wooding, 1956; Andersen et al., 1995; Díaz-García, 2013;), so that $E(A_\star) = \underline{\mu}_\star$ and $V(vec(A_\star)) = \nu I_{r_3}\otimes\Sigma$, where $\underline{\mu}_\star = \underline{\mu}_{\star R} + i\underline{\mu}_{\star I}$; note that imposing such a distribution for $A_\star$ is equivalent to assuming that $\begin{pmatrix} A_R \\ A_I \end{pmatrix}|\Sigma \sim MN\left(\begin{pmatrix} \underline{\mu}_{\star R} \\ \underline{\mu}_{\star I} \end{pmatrix}, \nu I_{r_3}, \begin{pmatrix} \frac{1}{2}\Sigma & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \frac{1}{2}\Sigma \end{pmatrix}\right)$ (Andersen et al., 1995),
- the parameter ν may be estimated or settled by the researcher; in the case of estimated ν we propose an inverse gamma distribution - $\nu \sim iG(\underline{s}_\nu, \underline{n}_\nu)$,
- the matrix normal distribution for the un-normalized cointegrating vectors at zero frequency: $B_1 \sim MN(0, \frac{1}{m_1}I_{r_1}, P_1)$, which results in the matrix angular central distribution for its orientation: $\beta_1 \sim MACG(P_1)$ (Chikuse, 1990; 2003); *via* matrix P_1 the researcher can incorporate prior knowledge about the cointegration space at zero frequency (see Koop et al. (2009) for details),
- the matrix normal distribution for the un-normalized cointegrating vectors at frequency π: $B_2 \sim MN(0, \frac{1}{m_2}I_{r_2}, P_2)$, so that $\beta_2 \sim MACG(P_2)$ (see the explanation for B_1),
- the complex matrix normal distribution for the un-normalized cointegrating vectors at frequency $\frac{\pi}{2}$: $B_\star \sim mCN(\mathbf{0}_{m_r\times r_3}, \frac{1}{m_3}I_{r_3}, P_\star)$, where $P_\star = P_{\star R} + iP_{\star I}$ is a Hermitian positive definite matrix ($P_\star = \bar{P}_\star'$), so for the real and imaginary parts of $B_\star$ we impose the matrix normal distribution of the form $\begin{pmatrix} B_R \\ B_I \end{pmatrix} \sim MN\left(0_{2m_3\times r_3}, \frac{1}{m_3}I_{r_3}, \frac{1}{2}\begin{pmatrix} P_{\star R} & -P_{\star I} \\ P_{\star I} & P_{\star R} \end{pmatrix}\right)$.

As regards to the prior distribution for $B_\star$ (see the final bullet point in the list above), Wróblewska (2020) has shown that such a distribution leads to the complex matrix angular central Gaussian distribution for the orientation part of the matrix $B_\star$, i.e. $\beta_\star \sim CMACG(P_\star)$. It is worth noting that, similarly to the *MACG* distribution in the real case, the *CMACG* distribution is invariant to multiplying $P_\star$ by some positive finite constant: $CMACG(P_\star) = CMACG(cP_\star),\ c > 0$. Again, similarly to the real case, one can introduce some prior knowledge about the cointegration space through $P_\star$, defining it as a weighted sum of the orthogonal projection matrix on a preferred cointegration space ($H_\star \bar{H}_\star'$) and its orthogonal complement ($H_{\star\perp} \bar{H}_{\star\perp}'$): $P_\star = H_\star \bar{H}_\star' + \tau H_{\star\perp} \bar{H}_{\star\perp}'$, where $\tau \in [0,1]$. For $\tau \in [0,1)$, the cointegration space spanned by $H_\star$ prevails, while $\tau = 1$ leads to the uniform distribution for the cointegration space. Finally, under $\tau = 0$ we obtain a degenerate (i.e. single-point) distribution, concentrated at $H_\star \bar{H}_\star'$. The τ parameter can be either fixed or estimated. In the latter case, similarly to Koop et al. (2009), one can specify for τ an inverse gamma prior distribution with such parameters that almost the entire probability mass is concentrated over [0,1]. Other choices are also admissible (such as the truncated inverse gamma or the uniform distribution).

The joint prior distribution is truncated by the non-explosive condition taking into account the appropriate number of unit roots at each frequency. The condition is imposed through the companion matrix of process (3) (e.g., Lütkepohl, 2005, pp. 15-16):

$$A = \begin{pmatrix} A_1 & A_2 & \dots & A_{k-1} & A_k \\ I_n & 0 & \dots & 0 & 0 \\ 0 & I_n & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_n & 0 \end{pmatrix}, \tag{6}$$

where I_n is an n-dimensional identity matrix, $A_1 = \Pi_1 + \Pi_2 + \Pi_4 + \Gamma_1$, $A_2 = \Pi_1 - \Pi_2 + \Pi_3 + \Gamma_2$, $A_3 = \Pi_1 + \Pi_2 - \Pi_4 + \Gamma_3$, $A_4 = I_n + \Pi_1 - \Pi_2 - \Pi_3 + \Gamma_4$, $A_i = \Gamma_i - \Gamma_{i-4}$ for $i = 5, 6, \dots$, $\Pi_1 = \alpha_1 \beta_1' \equiv A_1 B_1'$, $\Pi_2 = \alpha_2 \beta_2' \equiv A_1 B_1'$, $\Pi_3 = -2(\alpha_R \beta_R' + \alpha_I \beta_I') \equiv -2(A_R B_R' + A_I B_I')$, $\Pi_4 = 2(\alpha_I \beta_R' - \alpha_R \beta_I') \equiv 2(A_I B_R' - A_R B_I')$ and $\Gamma_i = 0$ for $i > k-4$. Note that the companion matrix related to a given VAR(k) process enables writing it in the form of VAR(1) (e.g., Lütkepohl, 2005, pp. 15-16). To ensure the non-explosiveness of VAR(k), all the eigenvalues of A need to remain within the unit circle, and the number of eigenvalues with unit module and corresponding to each of the frequencies should equal $n - r_j,\ j = 1, 2, 3$.

It is worth highlighting that in the prior structure presented above, the mutual independence of the cointegration spaces is assumed across the frequencies. Such an approach largely facilitates designing a MCMC procedure for the posterior inference, since the resulting full conditional posteriors of B_1, B_2 and $B_\star$ are of known form, thus enabling the Gibbs sampler, employed also by Koop et al. (2009); see the Appendix for details.

Arguably, the assumption of the prior independence between the spaces at different frequencies does not seem overly constraining, since the number as well as the form of the cointegrating vectors at these frequencies do not need to be related. Moreover, the assumption does not preclude testing the equality of the vectors across the frequencies, through Bayesian model comparison. Finally, it is also recognized in the frequentist (non-Bayesian) approach that such testing is valid without any prior knowledge about the existence of cointegration relations at other frequencies. We refer to Lee (1992) for further details.

The above-specified prior structure combined with the likelihood function leads to the joint posterior distribution with the following kernel:

$$\begin{aligned} p(\theta|y) \propto\ & \nu^{-\underline{n}_\nu - \frac{n}{2}[n(k-4)+l+r_1+r_2+2r_3]-1} \exp\left(-\frac{\underline{S}_\nu}{\nu}\right) \times \\ & \times |S|^{(q+q_s-n-1)/2} |\Sigma|^{-[q+n(k-4)+l+r_1+r_2+2r_3+T+n+1]/2} \times \\ & \times \exp\left\{-\frac{1}{2} tr\left[(A_s^{-1} + \Sigma^{-1})S\right]\right\} \exp\left\{-\frac{1}{2} tr\left[\Sigma^{-1} E'E\right]\right\} \times \\ & \times \exp\left\{-\frac{1}{2} tr\left[\frac{1}{\nu}\Sigma^{-1}\left(\Gamma - \underline{\mu}_\Gamma\right)' \underline{\Omega}_\Gamma^{-1} (\Gamma - \underline{\mu}_\Gamma)\right]\right\} \times \\ & \times \exp\left\{-\frac{1}{2} tr\left[\frac{2}{\nu}\Sigma^{-1}(A_\star - \underline{\mu}_\star)(A_\star - \underline{\mu}_\star)'\right]\right\} \times \\ & \times \exp\left\{-\frac{1}{2} tr\left[\frac{1}{\nu}\Sigma^{-1}(A_1 - \underline{\mu}_1)\underline{\Omega}_1^{-1}(A_1 - \underline{\mu}_1)'\right]\right\} \times \\ & \times \exp\left\{-\frac{1}{2} tr\left[\frac{1}{\nu}\Sigma^{-1}(A_2 - \underline{\mu}_2)\underline{\Omega}_2^{-1}(A_2 - \underline{\mu}_2)'\right]\right\} \times \\ & \times \exp\left\{-\frac{1}{2} tr\left[m_1 B_1' P_1^{-1} B_1 + m_2 B_2' P_2^{-1} B_2 + 2m_3 \tilde{B}_\star' P_\star^{-1} B_\star\right]\right\} I_{[0,1]}(|\lambda|), \end{aligned} \tag{7}$$

where $\theta = (\Sigma, \Gamma, A_1, A_2, A_\star, B_1, B_2, B_\star)$ collects all the model's parameters, $E = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)'$, l denotes the number of deterministic components gathered in $\tilde{D}_t$, $I_{[a,b]}(\cdot)$ is the indicator function of the interval $[a, b]$, and λ stands for the vector of the eigenvalues of the companion matrix. As already mentioned before, the pseudo-random sample from (8) may be obtained through the Gibbs sampler, similarly as in the models with cointegration at the zero frequency, following Koop et al. (2009). The set of full conditional posterior distributions and the sampling scheme are presented in the Appendix.

3. Point estimation of the cointegration space and the matrix of adjustment coefficients

Information about cointegration spaces obtained from the data may be summarized in the form of point estimates of these spaces along with some measure of the dispersion of these spaces' posterior distributions. Villani (2006) proposed employing the Frobenius (Hilbert-Schmidt) matrix norm for building the loss function used for the point estimation of real cointegration spaces. However, this approach can be adopted to estimate also complex spaces (e.g., Srivastava, 2000).

By employing the Frobenius matrix norm $\|A\|_F = (tr(\bar{A}'A))^{\frac{1}{2}}$ to the projection matrices we can build the loss function:

$$l(\beta, \tilde{\beta}) = \|\beta\bar{\beta}' - \tilde{\beta}\bar{\tilde{\beta}}'\|_F^2 = 2(r - tr(\beta\bar{\beta}'\tilde{\beta}\bar{\tilde{\beta}}')), \tag{8}$$

where r denotes the number of cointegrating vectors. It can be shown that this expected loss reaches its minimum in:

$$\hat{\beta} = \begin{pmatrix} \nu_1 & \nu_2 & \dots & \nu_r \end{pmatrix}, \tag{9}$$

where ν_i $(i = 1, 2, \dots, r)$ is the eigenvector of matrix $E(\beta\bar{\beta}')$, corresponding to its i-th largest eigenvalue (Chikuse, 2003; Villani, 2006). The numerical realization of $\hat{\beta}$ is obtained with the use of a pseudo-random sample from the posterior distribution of β, $\{\beta^{(s)},\ s = 1, 2, \dots, S\}$, and by approximating $E(\beta\bar{\beta}')$ with $\frac{1}{S}\sum_{s=1}^{S}\beta^{(s)}\bar{\beta}^{(s)'}$.

To measure the dispersion of the posterior distribution of each cointegration space, we follow Villani (2006) and use the projective Frobenius span variation:

$$\tau^2_{sp(\beta)} = \frac{r - \sum_{i=1}^{r}\lambda_i}{r(m-r)/m} \in [0, 1], \tag{10}$$

where λ_i is the i-th largest eigenvalue of $E(\beta\bar{\beta}')$. The measure $\tau^2_{sp(\beta)}$ reaches its minimum when the distribution is degenerate (i.e. deterministic, taking a single value with probability 1), whereas it hits the maximum for the uniform distribution over the (complex) Grassmann manifold, $\mathbb{G}_{r,m-r}$.

As mentioned in Section 2, in general, the cointegrating vectors and adjustment coefficients are identified up to an orthogonal transformation. Under only one cointegration relation, to ensure that the cointegrating vector is fully identified (by determining its orientation), it suffices to set the sign of any given coordinate of this vector. However, under more than only one cointegrating vector, as well as under complex vectors, such a restriction turns out to be insufficient. Although the estimation approach adopted in this work still yields unique point estimates of the cointegrating vectors, the non-identifiability is transferred onto the adjustment coefficients. To solve the non-identification problem in the general case one has to rotate all the draws of orthonormal cointegrating vectors $(\beta^{(s)})$ so that to minimize the distance between these vectors and the point estimate of their matrix. Additionally, the sign of one coordinate in each column of βs needs to be set. It is required that the adjustment coefficients be transformed as well, in accordance to the rotations of $\beta^{(s)}$. The algorithm proceeds as follows:

1. Based on $\{\beta^{(s)},\ s = 1, 2, \dots, S\}$, obtain the point estimate of the cointegrating vectors $\hat{\beta}$.
2. Calculate and store the distances $d^{(s)}$ between $\hat{\beta}$ and each $\beta^{(s)}$, $s = 1, 2, \dots, S$.
3. For each $s \in \{1, 2, \dots, S\}$, draw the orthonormal matrix $O^{(s)}$ of order r, and then calculate $\tilde{\beta}^{(s)} = \beta^{(s)}O^{(s)}$. Check the sign restriction: If not satisfied, then create matrix $H^{(s)} = diag(\pm 1)$, which can be used to ensure relevant signs: $\tilde{\beta}^{(s)} = \beta^{(s)}O^{(s)}H^{(s)}$.
4. Calculate the new distance $\tilde{d}^{(s)}$: If smaller than $d^{(s)}$, then calculate $\tilde{\alpha}^{(s)} = \alpha^{(s)}(O^{(s)})^{-1}$ (or $\tilde{\alpha}^{(s)} = \alpha^{(s)}H^{(s)}(O^{(s)})^{-1}$, if the sign correction was necessary) and store the new draws.
5. Repeat steps 2-4 until the stop condition is reached.

Note that by the performed transformation neither the cointegration space nor the $\Pi^{(s)}$ matrices are changed, i.e. the point estimate of the cointegrating space stays the same and $\tilde{\alpha}^{(s)}\tilde{\beta}^{(s)'} = \alpha^{(s)}\bar{\beta}^{(s)'}$.

The distance between cointegrating matrices can be measured by the angle obtained as $d^{(s)} = \arccos(tr(\beta^{(s)'}\hat{\beta})/r)$ for the real matrices. In the complex case, one can use the same measure but applied to matrices created by stacking the real and imaginary parts, i.e. for $\beta^{(s)} = \begin{pmatrix}\beta_R^{(s)}\\ \beta_I^{(s)}\end{pmatrix}$ and $\hat{\beta} = \begin{pmatrix}\hat{\beta}_R\\ \hat{\beta}_I\end{pmatrix}$.

4. Empirical illustration

In the empirical analysis, we consider four-dimensional models for the Polish GDP (in constant 2010 prices), the annualized inflation rate based on the consumer price index (CPI, 2015 = 100), the real broad monetary aggregate M3, and the spread between the long- and short-term interest rates approximated by the difference between the 10-Year Bond Yield and 3-month Warsaw Inter Bank Offer Rate (WIBOR). All the variables (except for the interest rates spread) are modeled in logs. Such a specification, usually referred to as the P^* model, was introduced by Hallman et al. (1991) and relates the long-run equilibrium price level to the current value of the monetary aggregate M2, the long-run equilibrium value of money velocity, and the current value of potential real GNP. The model has been employed in numerous empirical studies, performed either for prices and the nominal money stock (e.g., Tödter and Reimers, 1994; Herwartz and Reimers, 2006) or for inflation and

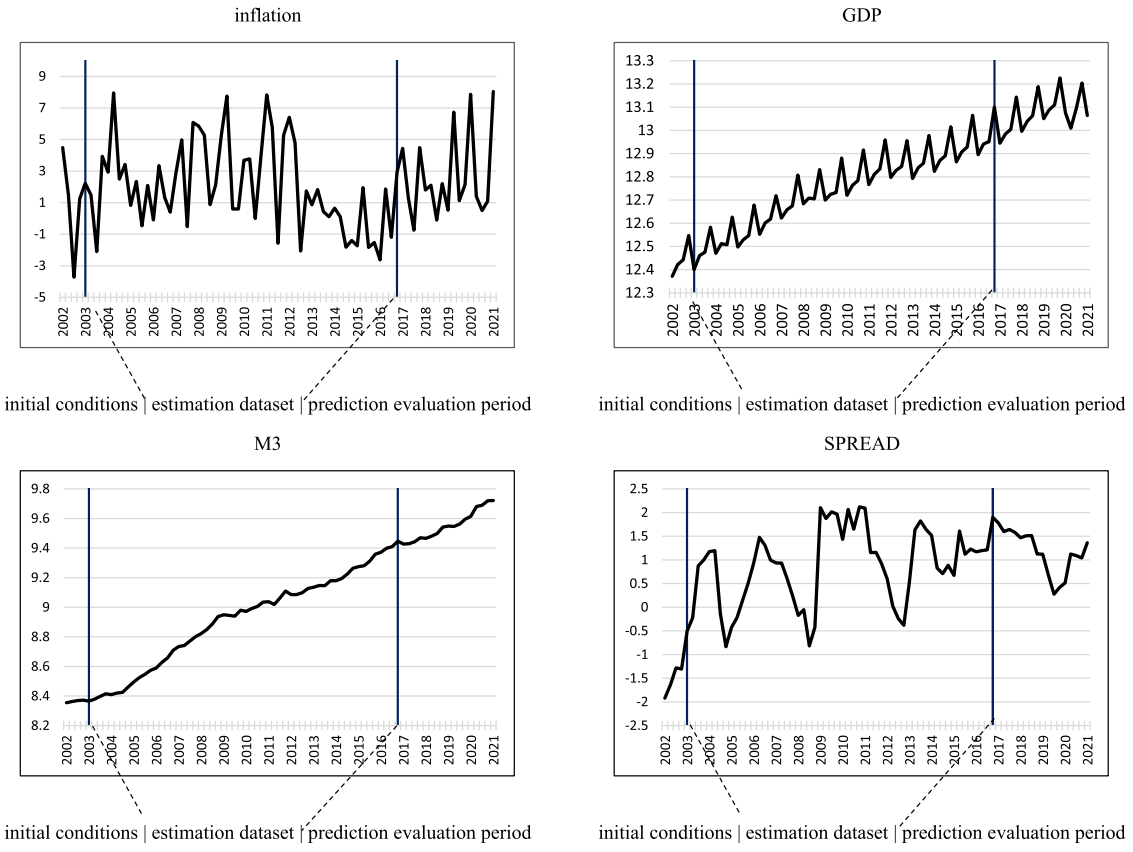


Fig. 1. The analyzed data: initial conditions (2002Q1-2003Q1) | estimation dataset (2003Q2-2016Q4) | prediction evaluation period (2017Q1-2022Q2).

the real money aggregate (e.g., Wesche, 1997; Rudebusch and Svensson, 2002; Gerlach and Svensson, 2003). For (seasonal) quarterly data representing the Polish economy, a specification similar to ours was analyzed by Kotłowski (2005).

According to the discussion presented by Gerlach and Svensson (2003), in the model built for the real variables, two cointegrating relations can be expected: one corresponding to the long-run money demand equation, and the other describing the covariance stationarity of the interest rates spread.

The data on the GDP, CPI and monetary aggregate have been retrieved from Federal Reserve Bank of St. Louis (https://fred.stlouisfed.org), while the long- and short-term interest rates have been obtained from Investing.com and Stooq.pl services, respectively.

The series under study covers quarterly observations over the period 2002Q1 through 2022Q2. Data up to 2016Q4 is used for the in-sample inference (model comparison and estimation of cointegration spaces) and also as the initial sample for a forecasting exercise over the following 22 quarters. Figure 1 depicts the analyzed time series, with markings of the forecast evaluation period and initial conditions. Apparently, all time series under consideration display more or less discernible trending behavior. Additionally, two of the analyzed series (GDP and inflation) manifest clear seasonal patterns, also seen in the unit-root-transformed series (see Figure 2). The transformed data, denoted as $y_t^{(1)}$, $y_t^{(2)}$ and $y_t^{(3)}$, have been obtained through the approach proposed by Engle et al. (1993), with the filtration performed at frequencies of zero, π and $\frac{\pi}{2}$, respectively. The seasonality of the inflation and GDP can also be confirmed by visual inspection of their autocorrelation and partial autocorrelation functions, displaying conspicuous spikes at the seasonal lags (see Figure 3). Moreover, the patterns of the ACFs of the GDP and M3 confirm the trending behavior of these two series, whereas the ACFs of the inflation and spread may indicate covariance stationarity. These hypotheses can be formally checked by means of Bayesian model comparison.

To fully define the Bayesian seasonally cointegrated VAR model, we impose the following priors:

$S \sim W(0.02I_4, 5)$, $\Sigma \sim iW(S, 6)$, $\Gamma|\Sigma \sim MN\big(\mathbf{0}, \Sigma, \nu I_{4+l}\big)$, where l denotes the number of dummies outside cointegration spaces, $A_1|\Sigma \sim MN(\mathbf{0}_{n\times r_1}, \nu I_{r_1}, \Sigma)$, $B_1 \sim MN(0, \frac{1}{m_1} I_{r_1}, 0.1I_4)$, $A_2|\Sigma \sim MN(\mathbf{0}_{n\times r_2}, \nu I_{r_2}, \Sigma)$, $B_2 \sim MN(0, \frac{1}{m_2} I_{r_2}, 0.1I_4)$, $A_\star|\Sigma \sim mCN(\mathbf{0}_{n\times r_3}, \nu I_{r_3}, \Sigma)$, $B_\star \sim mCN(\mathbf{0}_{m_r\times r_3}, \frac{1}{m_3} I_{r_3}, 0.1I_4)$, $\nu \sim iG(2, 3)$. The prior variance of the matrices B_i, $i = 1, 2, \star$, is scaled by 0.1 for numerical reasons, to increase the number of draws that satisfy the non-explosive condition. Note that it does not change the prior assumptions of the cointegration space.

To investigate the nature of seasonal fluctuations in the data, we compare alternative model specifications differing in two respects: firstly, the number of cointegrating relations ($r_j \in \{0, 1, 2, 3, 4\}$ for $j = 1, 2, 3$) at the frequencies of zero ($j = 1$),

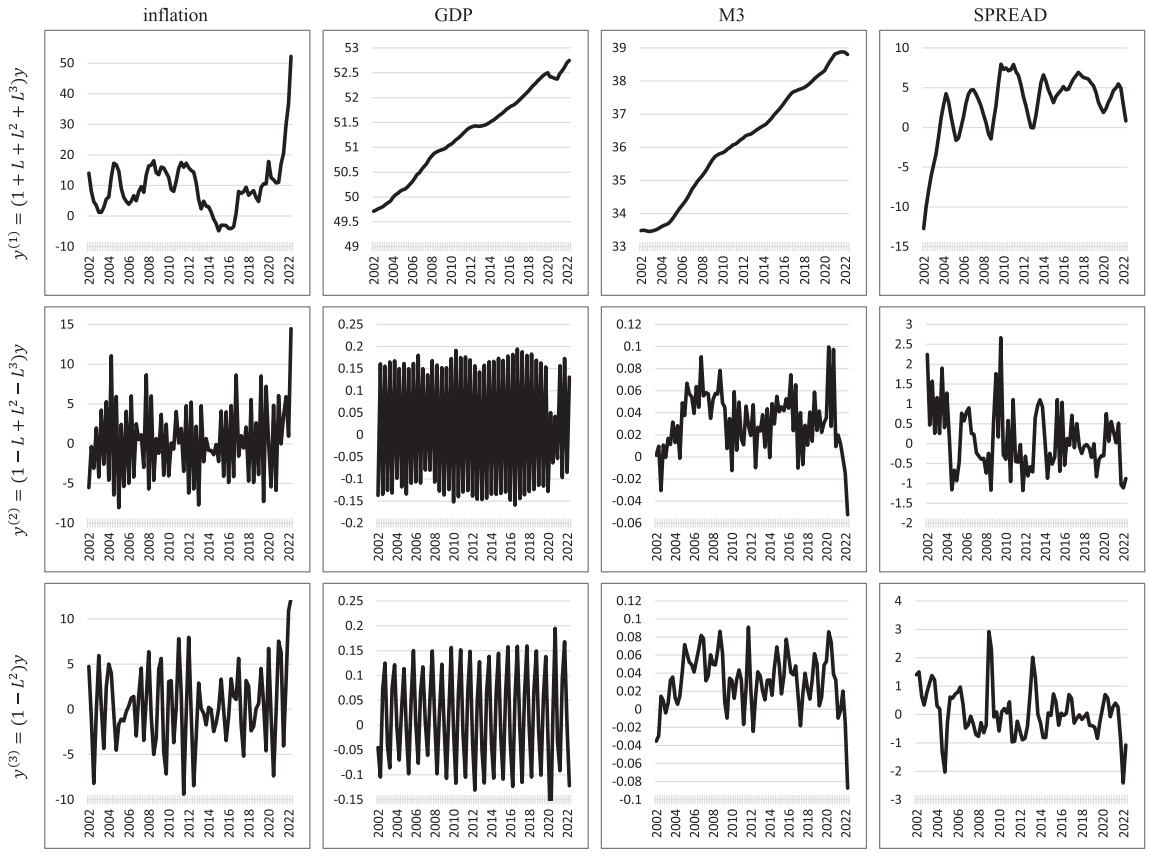


Fig. 2. The unit root transformations of the analyzed data: $y^{(1)} = (1 + L + L^2 + L^3)y$, $y^{(2)} = (1 - L + L^2 - L^3)y$, $y^{(3)} = (1 - L^2)y$.

π ($j = 2$) and $\frac{\pi}{2}$ ($j = 3$), and secondly, the inclusion of seasonal dummies, with $s = 1$ ($s = 0$) denoting models with (without) such variables. Additionally, we consider models in which either the inflation or interest rates spread (but not both) are assumed to be stationary at frequency 0. In such specifications, we test the hypothesis that some cointegrating vectors are known, i.e. $\beta = (b, \ \psi)$ with a known matrix b and estimated ψ, so that $sp(b) \subseteq sp(\beta)$. To test the stationarity, b is assumed to be such a unit vector that it picks the tested variable; see, e.g., Johansen (1995), Juselius (2006) for the more elaborate discussion, and Strachan and van Dijk (2007), Wróblewska (2009) for an analysis within the Bayesian framework. Each of the considered specifications features five lags in their VAR representation and an unrestricted constant in the VEC specification. Note that for the sake of our analysis it is required for quarterly data to allow for at least $k = 5$ lags in the underlying VAR structures. Setting only $k = 4$ lags (as could be conceived of for quarterly observations) would cancel out the short-term dynamics in Eq. (3), which would appear unwanted. On the other hand, a higher number of lags could be considered. However, bearing in mind a rather low sample size of the modeled series ($T = 55$), and also to avoid highly time-consuming computations for another couple hundred of model specifications, we decided to restrict the study to the five-lagged VARs only.

After excluding non-possible feature combinations and leaving in the set one representation of each of the observationally equivalent models, we are left with 637 pairwise different models. For each of them, we assume equal prior probabilities, i.e. $p(M_{s,r_1,r_2,r_3,c}) = \frac{1}{637} \approx 0.0016$, where $c \in \{0, 1, 2\}$, with $c = 0$ denoting a model without additional restrictions imposed on the cointegration space at frequency 0, $c = 1$ – a model with a stationary interest rates spread at frequency 0, and finally, $c = 2$ – a model with a stationary inflation rate at frequency 0. Note that imposing the uniform prior on the models' space does not lead to a uniform prior distribution for the model features (Strachan and van Dijk, 2007).

As noted by Geweke (2001,2005), Geweke and Amisano (2010,2011), Koop et al. (2011), to compare models one can actually employ one-step-ahead predictive likelihoods, typically associated with evaluation of forecasts' accuracy. According to Theorem 2.6.1 in Geweke (2005, p. 66), the predictive likelihood updates the marginal likelihood and produces the new marginal likelihood (see also Geisel, 1973), so that the final marginal likelihood can be decomposed into the product of consecutive one-step-ahead predictive likelihoods. Therefore, within the Bayesian setting, cumulative (one-step-ahead) predictive likelihoods (henceforth, CLPL) obtained for alternative models over a given range of data can formally be used for comparing models' conformity with modeled data. Such an approach is commonly advocated in the case of models for which calculation of the marginal data density poses somewhat a challenge. The works cited above provide the details.

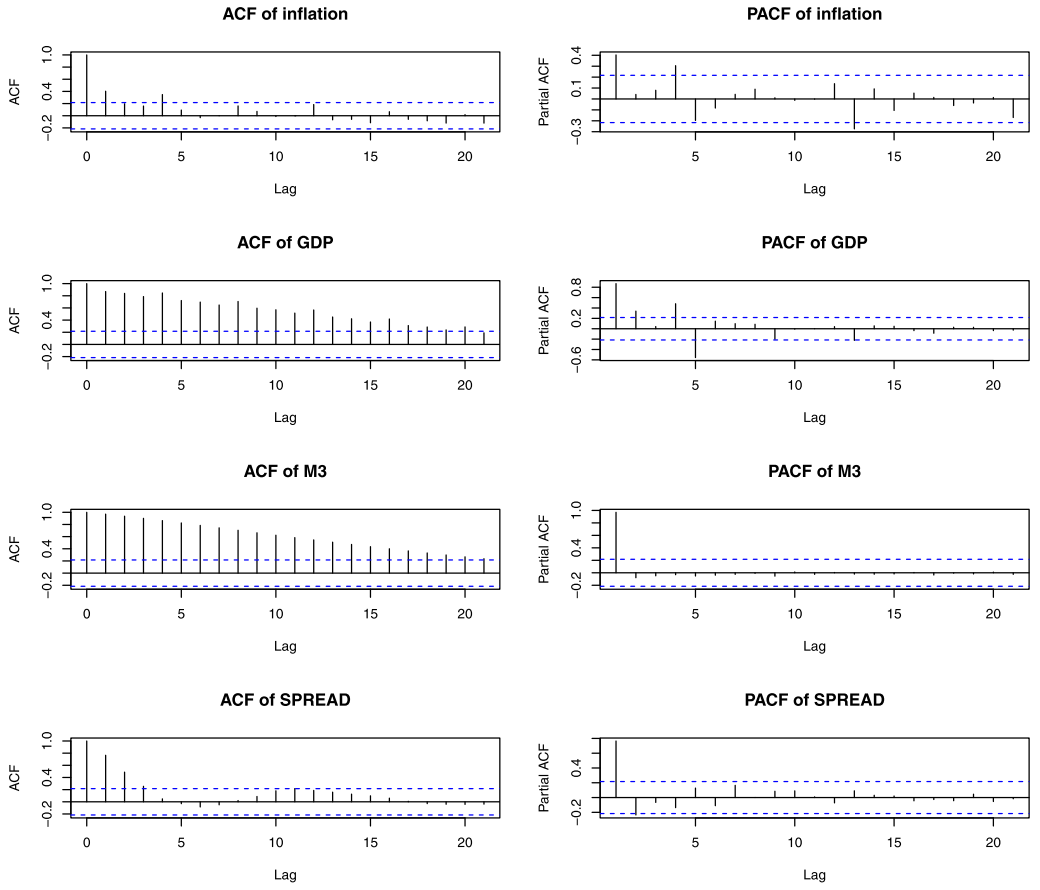


Fig. 3. The autocorrelation and partial autocorrelation functions.

Table 1
Decimal logarithm of cumulative predictive likelihoods (CLPL), along with cumulative log predictive Bayes factors (CLPBF) for models with $CLPBF_{1i} < 3$, calculated in favor of the best specification

Rank (i)	s	r_1	r_2	r_3	c	$CLPL_i$	$CLPBF_{1i}$
1	0	0	2	1	0	0.715	0
2	0	0	3	0	0	0.446	0.270
3	0	2	4	0	0	-0.277	0.992
4	0	1	0	2	0	-0.826	1.542
5	1	1	3	1	0	-0.964	1.680
6	1	2	0	3	2	-1.117	1.833
7	1	1	1	2	0	-1.185	1.900
8	0	1	4	1	0	-1.802	2.517
9	1	1	1	0	0	-2.013	2.728
10	1	2	2	1	0	-2.037	2.753
11	0	1	1	0	0	-2.044	2.760

Calculations based on 1000 accepted draws from the posterior distribution of the models built for the each of 54 considered observations. For the first observation we set 10000 burn-in draws. Upon including each of the subsequent data points, each time the MCMC sampler is initiated at posterior estimates obtained in the previous run, and only 1000 burn-in draws are generated.

Table 1 displays models with highest values of cumulative log predictive likelihood (CLPL) for the last 54 modeled observations, i.e. covering the period 2003Q3-2016Q4. Cumulative (decimal) log predictive Bayes factors ($CLPBF_{1i}$ with positive values in favor of the best model) are also included. For the sake of exposition, only models with $CLPBF_{1i}$ no greater than 3 are presented, which gives 11 specifications.

The best-performing models include the ones with as well as without seasonal deterministic constants. In general, however, the top part of the ranking is dominated by some specifications that exclude seasonal dummies. Moreover, the total posterior probability of all the models precluding seasonal intercepts is as high as approximately 0.969 (see Table 2), indicating an overpowering evidence in favor of these specifications.

Table 2
Posterior probabilities of the models' features

s	$p(\cdot\|y)$	r_1	$p(\cdot\|y)$	r_2	$p(\cdot\|y)$	r_3	$p(\cdot\|y)$	c	$p(\cdot\|y)$
0	0.969	0	0.890	0	0.025	0	0.373	0	0.991
1	0.031	1	0.040	1	0.009	1	0.594	1	0.000
		2	0.070	2	0.580	2	0.024	2	0.009
		3	0.000	3	0.323	3	0.009		
		4	0.000	4	0.061	4	0.000		

Calculations based on 1000 accepted draws from the posterior distribution of the models built for the each of 54 considered observations. For the first observation we set 10000 burn-in draws. Upon including each of the subsequent data points, each time the MCMC sampler is initiated at posterior estimates obtained in the previous run, and only 1000 burn-in draws are generated.

The best-performing model with the overidentifying restriction of the spread's stationarity at frequency 0, is ranked 6, with $CLPBF_{1i} = 1.833$ (see the results for $M_{1,2,0,3,2}$ in Table 1). On the whole, the models assuming long-run stationarity of the interest rates' spread collect the posterior probability of only 0.006. As regards the models hypothesizing the long-run stationarity of inflation, none of them has made its way to Table 1 (incidentally, the total posterior probability of these specifications is close to zero). Almost all of the top 11 models, ranked in Table 1, feature seasonal cointegration or integration both at the annual as well as bi-annual frequency, thereby providing strong empirical evidence for seasonal trends in the analyzed series. Generally, the data support models with two or three cointegrating vectors at the bi-annual frequency and with none or one vector at the annual frequency. Interestingly, the models with cointegration at the zero frequency are only marginally supported by the data.

Below we supplement the above considerations on the nature of the analyzed data, with inference about the cointegration spaces and the adjustment coefficients in the model ranked first, i.e., $M_{0,0,2,1,0}$. The cointegrating vectors' point estimates are obtained through the approach discussed in Section 3. For the adjustment coefficients posterior medians as well as 0.16 and 0.84 quantiles (in brackets) are provided. The elements for which the value of zero does not fall within the corresponding 68% credible intervals are marked in bold. Finally, since the model at hand does not feature cointegration at the zero frequency, we present the results only for the other two frequencies:

- $r_2 = 2$,

$$\hat{\beta}_2 = \begin{pmatrix} 0.005 & 0.051 \\ 0.033 & -0.849 \\ -0.031 & -0.526 \\ -0.999 & -0.012 \end{pmatrix}, \tau^2_{sp(\beta_2)} = 0.653, \tag{11}$$

$$\hat{\alpha}_2 = \begin{pmatrix} -0.449 & 0.070 \\ -0.001 & 0.000 \\ -0.004 & -0.001 \\ \mathbf{-0.358} & -0.001 \end{pmatrix} \begin{matrix} (-0.959,\ 0.000) & (-0.393,\ 0.592) \\ (-0.004,\ 0.003) & (-0.004,\ 0.003) \\ (-0.012,\ 0.003) & (-0.008,\ 0.005) \\ (-0.532,\ -0.187) & (-0.212,\ 0.207) \end{matrix}, \tag{12}$$

- $r_3 = 1$,

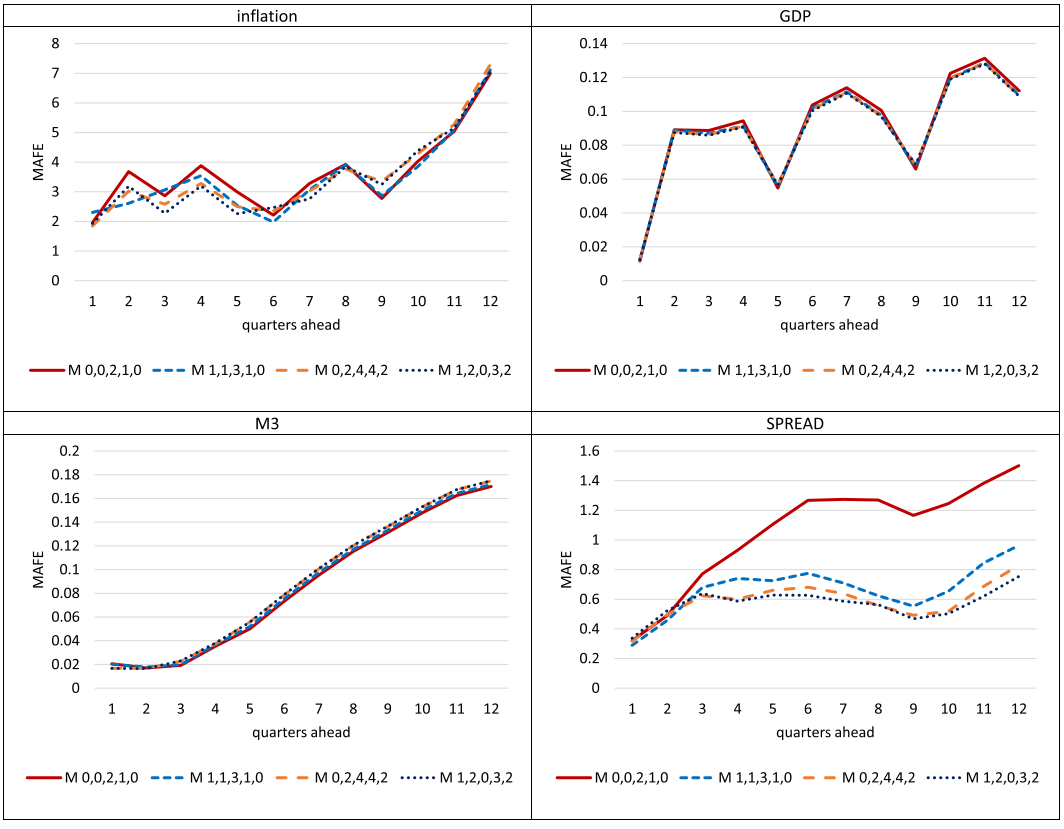
$$\hat{\beta}_* = \begin{pmatrix} 0.060 \\ 0.033 \\ 0.009 \\ -0.996 \end{pmatrix} + i \begin{pmatrix} -0.021 \\ -0.043 \\ 0.040 \\ 0.000 \end{pmatrix}, \tau^2_{sp(\beta_*)} = 0.403, \tag{13}$$

$$\hat{\alpha}_* = \begin{pmatrix} 0.110 \\ -0.003 \\ 0.002 \\ -0.023 \end{pmatrix} \begin{matrix} (-0.198,\ 0.422) \\ (-0.006,\ 0.002) \\ (-0.004,\ 0.007) \\ (-0.391,\ 0.191) \end{matrix} + i \begin{pmatrix} -0.164 \\ -0.001 \\ 0.002 \\ 0.412 \end{pmatrix} \begin{matrix} (-0.531,\ 0.338) \\ (-0.004,\ 0.002) \\ (-0.003,\ 0.006) \\ (-0.189,\ 0.536) \end{matrix}, \tag{14}$$

As implied by the values of $\tau^2_{sp(\beta_2)}$ and $\tau^2_{sp(\beta_*)}$, the posterior distributions of the cointegrating spaces at the two seasonal frequencies are quite diffused. As for the short-term adjustments, none of the considered variables seems to adjust significantly to the annual cointegration relation, as zero falls within the 68% credible intervals. However, the spread adjusts to the first cointegrating relation at the bi-annual frequency. Finally, since the estimates of the first cointegrating vector at the bi-annual frequency, and the real part of the cointegrating vector at the annual frequency, resemble the identity vector $(0, 0, 0, 1)'$, the stationarity of the spread at these (seasonal) frequencies is clearly implied.

Additionally, we provide the point estimates (posterior medians) of the matrices Π_i, $i = 2, 3, 4$:

$$\hat{\Pi}_2 = \begin{pmatrix} 0.049 & \mathbf{-0.110} & \mathbf{0.018} & \mathbf{0.435} \\ 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.001 & 0.001 & 0.004 \\ -0.004 & \mathbf{-0.008} & 0.009 & \mathbf{0.333} \end{pmatrix},$$



Calculations based on 10000 accepted draws from the predictive distribution. For the first iteration we set 100000 *burn-in* draws. Upon including each of the subsequent data points, each time the MCMC sampler is initiated at posterior estimates obtained in the previous run, and only 10000 *burn-in* draws are generated.

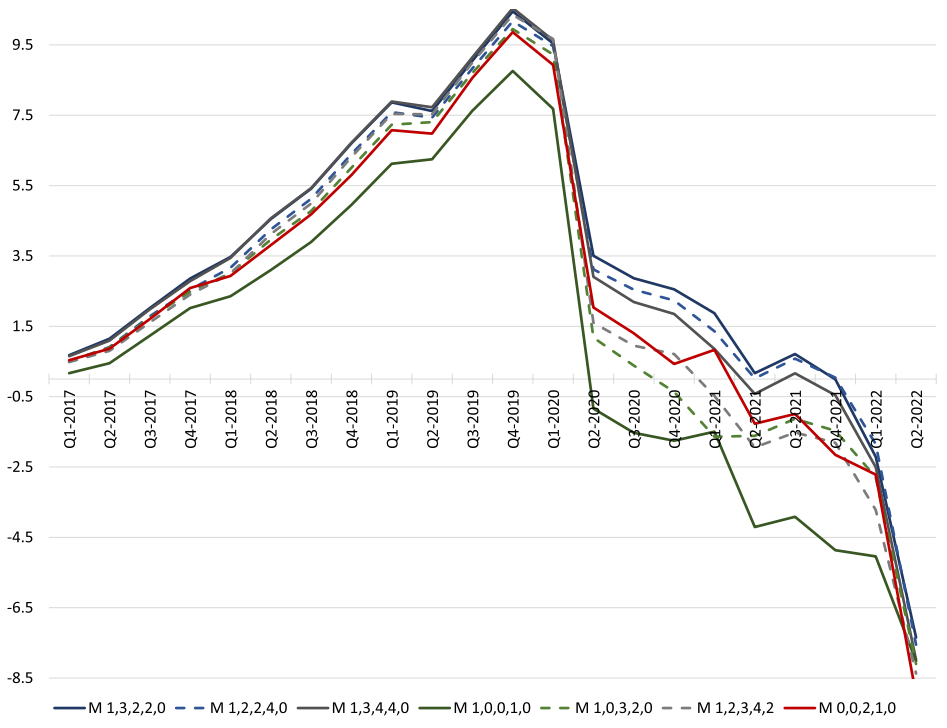
Fig. 4. Mean absolute forecast errors.

Table 3
Decimal logarithm of cumulative predictive likelihoods (CLPL), along with cumulative log predictive Bayes factors (CLPBF) for chosen models, calculated in favor of the best specification

Rank (i)	s	r_1	r_2	r_3	c	$CLPL_i$	$CLPBF_{1i}$
1	1	3	2	2	0	-7.347	0.000
2	1	2	2	4	0	-7.553	0.205
3	1	3	4	4	0	-7.985	0.637
4	1	0	0	1	0	-8.024	0.677
5	1	0	3	2	0	-8.101	0.753
6	1	3	3	3	0	-8.207	0.859
7	1	2	3	4	2	-8.374	1.026
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	0	0	2	1	0	-8.980	1.632

$$\hat{\Pi}_3 = \begin{pmatrix} 0.041 & \mathbf{0.034} & \mathbf{0.010} & \mathbf{-0.554} \\ 0.000 & 0.000 & 0.000 & 0.003 \\ 0.000 & 0.000 & 0.001 & -0.001 \\ -0.048 & \mathbf{-0.041} & 0.013 & \mathbf{0.647} \end{pmatrix},$$

$$\hat{\Pi}_4 = \begin{pmatrix} -0.012 & \mathbf{0.005} & \mathbf{0.030} & \mathbf{0.061} \\ 0.000 & 0.001 & 0.000 & -0.006 \\ 0.000 & 0.000 & 0.000 & 0.006 \\ -0.015 & \mathbf{0.017} & \mathbf{-0.028} & \mathbf{0.476} \end{pmatrix}. \tag{15}$$



Calculations based on 10000 accepted draws from the predictive distribution. For the first iteration we set 100000 *burn-in* draws. Upon including each of the subsequent data points, each time the MCMC sampler is initiated at posterior estimates obtained in the previous run, and only 10000 *burn-in* draws are generated.

Fig. 5. Decimal logarithm of cumulative predictive likelihoods (CLPL).

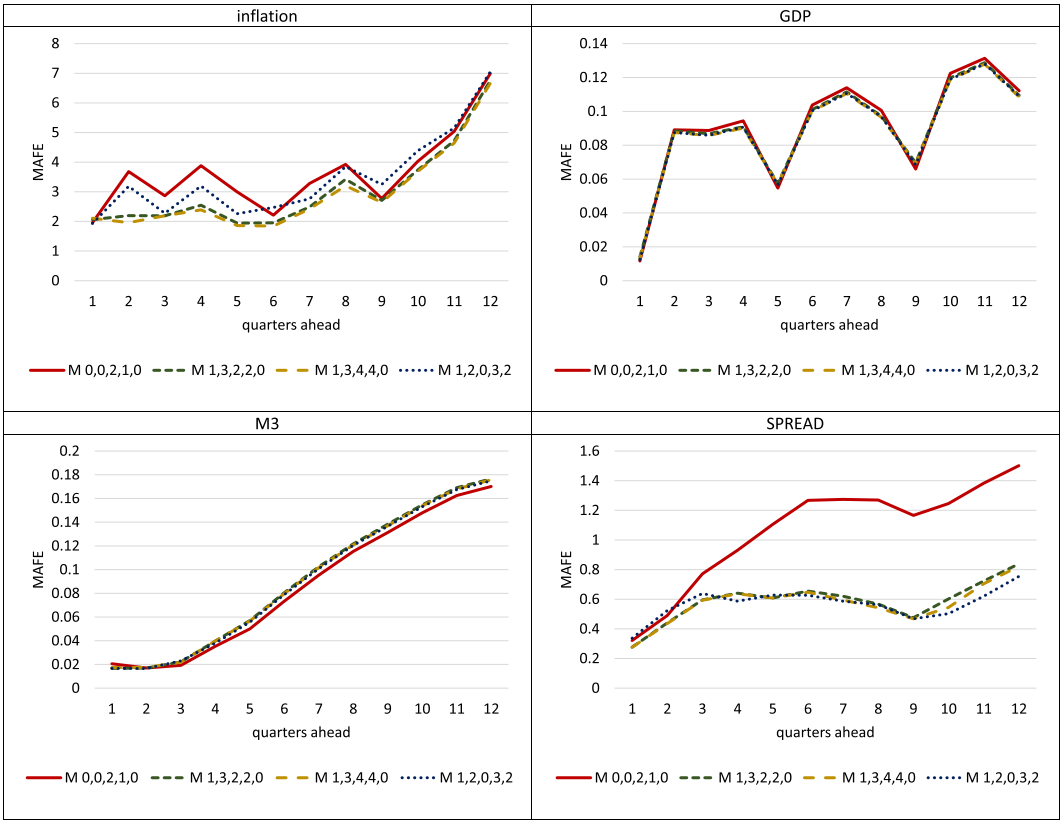
In all the matrices above, there are some elements (in bold) for which the value of zero does not fall within the corresponding 68% credible intervals (unreported, but available upon request). The outcome indicates that seasonal cointegrating relations should be incorporated in the model, as they are key for capturing the dynamics of the inflation and spread, even though the results point to no significant adjustment (measured by the point estimates of αs) to the cointegration relations.

In what follows, we examine the forecasting performance of the following 4 arbitrarily chosen specifications:

- the (winning) model with seasonal cointegration and without cointegration at the zero frequency: $M_{0,0,2,1,0}$,
- the best model featuring cointegration both at the zero and seasonal frequencies (ranked 5^{th}): $M_{1,1,3,1,0}$,
- the best model assuming the stationarity of the spread (ranked 6^{th}): $M_{0,2,0,3,2}$,
- the best model assuming the stationarity of the entire VAR system at the seasonal frequencies (ranked 38^{th}): $M_{0,2,4,4,2}$.

The ex-post prediction evaluation period ranges from 2017Q1 to 2022Q2 (22 quarters). In each of the 4 models and for each of the ex-post prediction data points, we produce 1- to 12-quarter-ahead forecasts, starting with the estimation sample covering the period 2002Q1 - 2016Q4. Then, the estimation sample and the posterior distribution are updated (one quarter at a time), and the procedure is reiterated until 2019Q2, so that for each of the examined models we obtain 11 forecast paths. For the point predictions we take the medians of the predictive distributions. To compare the forecast accuracy of the models, we employ mean absolute forecast errors (MAFE), computed at each of the considered horizons, with their values displayed in Figure 4. The forecast accuracy obtained for inflation, GDP and M3 is similar across all the models. However, regarding the forecast accuracy of the spread, the model $M_{0,0,2,1,0}$ performs the worst. Since this is the only model without cointegrating relations at the zero frequency, it can be concluded that allowing for long-run relationships improves the forecast accuracy of this variable. Moreover, the forecasts obtained in the models additionally assuming stationarity emerge the most accurate.

In general, surprisingly, the point forecasts produced by the model that proved the best within the sample, i.e. $M_{0,0,2,1,0}$, turned out worse than the ones obtained in the other specifications (although only with respect to the spread). As counter-intuitive as it may appear, a somewhat similar outcome has been reported previously by Pajor and Wróblewska (2022), who reveal that the point forecasts accuracy measure within the Bayesian framework should not be employed to pinpoint the 'correct' model specification (i.e. the one featuring the highest posterior probability). For this reason, we extend the analysis here by including also density forecasts assessment, by means of the cumulative log predictive likelihoods (CLPL;



Calculations based on 10000 accepted draws from the predictive distribution. For the first iteration we set 100000 *burn-in* draws. Upon including each of the subsequent data points, each time the MCMC sampler is initiated at posterior estimates obtained in the previous run, and only 10000 *burn-in* draws are generated.

Fig. 6. Mean absolute forecast errors.

the decimal logarithm is applied) and cumulative log predictive Bayes factors calculated over the forecast evaluation period, i.e. 2017Q1 - 2022Q2 (see Table 3 and Figure 5), and select two superior models (in terms of CLPL) to additionally include in the forecast accuracy comparison (still, with respect to MAFE, for the sake of consistency with the results discussed earlier; see Figure 6). With respect to the cumulative log predictive likelihood, the $M_{0,0,2,1,0}$ model is ranked 17^{th} and is 1.6 times worse than the winning specification, $M_{1,3,2,2,0}$; see Table 3. Interestingly, as opposed to the 'in-sample' comparison (see Table 1), among the models occupying the first seven places there are only those with seasonal dummies and, generally, the cointegration at the zero frequency is supported by the data.

According to Figure 5, all the way before the final data point of 2022Q2, the cumulative predictive likelihood of model $M_{1,0,0,1,0}$ (ranked fourth) was systematically below the CLPLs of all the remaining specifications presented in the figure (including even the one that ranked only 17^{th}, i.e. $M_{0,0,2,1,0}$). Only at this final date of 2022Q2 was the ranking of the models distorted, which can be attributed to abrupt data movements in the final two quarters, i.e. 2022Q1-Q2, particularly conspicuous in the unit-root-transformed series; see Figure 2.

Finally, the MAFE-based forecast accuracy of the winning specification, $M_{1,3,2,2,0}$, is much higher than the model that ranked first in the period 2003Q3-2016Q4.

5. Conclusions

In this paper, a Bayesian seasonally cointegrated vector error correction model for quarterly data was introduced. To that end, we defined relevant prior distributions and derived full conditional posteriors that underlie the Gibbs sampler, through which the joint posterior distribution can be sampled from. The point estimation of the cointegration spaces at considered frequencies was also discussed.

The Bayesian methodology allows for a formal model comparison even in a large set of specifications that may differ by several features at the same time, which proved conducive to our framework. Alternative models' relevance in view of given

data is assessed probabilistically, typically by pairs, and not by a sequence of tests. Such a procedure facilitates the selection of the best model specification (and also a model pooling approach to inference, if opted for), which is crucial for further stages of empirical analyses (e.g., Abeysinghe (1994) and Arteche (2007), but also the empirical analysis in this paper).

As pointed out in the Introduction, disregarding seasonal cointegration and handling seasonality through only dummy variables entails a risk of model misspecification, which can affect the results of both the in- and out-of-sample performance. The relevance of the methodology developed in the present work was illustrated by an analysis of four-dimensional time series pertaining to the Polish economy. The in-sample model comparison provides clear evidence for cointegration at the bi-annual and annual frequencies, indicating that seasonally-cointegrated VEC models should indeed be considered for the analysed data. In turn, for the sake of the out-of-sample performance the models appear to benefit not only from the seasonal cointegration, but additionally also from the zero-frequency cointegration and seasonal dummies, as well.

Declaration of Competing Interest

I declare no conflict of interest.

Acknowledgments

I acknowledge financial support from a subsidy granted to Krakow University of Economics - Project nr 080/EIE/2022/POT. I am truly grateful to the Editor, two anonymous Reviewers, as well as Łukasz Kwiatkowski and Peter Winker for valuable comments and suggestions, which helped me to improve the paper.

Appendix A. Full conditional posterior distributions and the sampling scheme

We introduce the following matrix form of model (3):

$$Z_0 = Z_1\beta_1\alpha_1' + Z_2\beta_2\alpha_2' + Z_3\bar{\beta}_\star\alpha_\star' + \bar{Z}_3\beta_\star\bar{\alpha}_\star' + Z_4\Gamma + E, \tag{A.1}$$

where $Z_0 = \left(\Delta_4 y_1 \quad \Delta_4 y_2 \quad \dots \quad \Delta_4 y_T\right)'$, $Z_1 = \left(\tilde{y}_1^{(1)} \quad \tilde{y}_2^{(1)} \quad \dots \quad \tilde{y}_T^{(1)}\right)'$, $Z_2 = \left(\tilde{y}_1^{(2)} \quad \tilde{y}_2^{(2)} \quad \dots \quad \tilde{y}_T^{(2)}\right)'$, $Z_3 = \left(\tilde{y}_1^{(3)} \quad \tilde{y}_2^{(3)} \quad \dots \quad \tilde{y}_T^{(3)}\right)' = -Z_{32} - iZ_{31}$, $Z_{31} = \left(\tilde{y}_1^{(31)} \quad \tilde{y}_2^{(31)} \quad \dots \quad \tilde{y}_T^{(31)}\right)'$, $Z_{32} = \left(\tilde{y}_1^{(32)} \quad \tilde{y}_2^{(32)} \quad \dots \quad \tilde{y}_T^{(32)}\right)'$, $Z_4 = \left(z_1 \quad z_2 \quad \dots \quad z_T\right)'$, $z_t' = \left(\Delta_4 y_{t-1}' \quad \Delta_4 y_{t-2}' \quad \dots \quad \Delta_4 y_{t-k+4}' \quad \tilde{D}_t\right)$, $\Gamma = \left(\Gamma_1 \quad \Gamma_2 \quad \dots \quad \Gamma_{k-4} \quad \tilde{\Phi}\right)'$, $E = \left(\varepsilon_1 \quad \varepsilon_2 \quad \dots \quad \varepsilon_T\right)'$.

From equation (8), the following set of full conditional posterior distributions can be derived:

- the inverse Wishart distribution for Σ (i.e. the covariance matrix of errors):

$$p(\Sigma|\cdot, y) = f_{iW}\left(\overline{S}, q + n(k-4) + l + r_1 + r_2 + 2r_3 + T\right), \tag{A.2}$$

where

$$\overline{S} = S + \frac{1}{\nu}\Big[(\Gamma - \underline{\mu}_\Gamma)'\underline{\Omega}_\Gamma^{-1}(\Gamma - \underline{\mu}_\Gamma) + 2(A_\star - \underline{\mu}_\star)(A_\star - \underline{\mu}_\star)' + \\ + (A_1 - \underline{\mu}_1)\underline{\Omega}_1^{-1}(A_1 - \underline{\mu}_1)' + (A_2 - \underline{\mu}_2)\underline{\Omega}_2^{-1}(A_2 - \underline{\mu}_2)'\Big] + E'E,$$

- the matrix normal distribution for Γ:

$$p(\Gamma|\cdot, y) = f_{MN}\left(\overline{\mu}_\Gamma, \Sigma, \overline{\Omega}_\Gamma\right), \tag{A.3}$$

where $\overline{\Omega}_\Gamma = (\frac{1}{\nu}\underline{\Omega}_\Gamma^{-1} + Z_4'Z_4)^{-1}$ and $\overline{\mu}_\Gamma = \overline{\Omega}_\Gamma\left[\frac{1}{\nu}\underline{\Omega}_\Gamma^{-1}\underline{\mu}_\Gamma + Z_4'\left(Z_0 - Z_1B_1A_1' - Z_2B_2A_2' - 2Re(Z_3\bar{B}_\star A_\star')\right)\right]$,

- the matrix normal distribution for A_1:

$$p(A_1|\cdot, y) = f_{MN}\left(\overline{\mu}_1, \overline{\Omega}_1, \Sigma\right), \tag{A.4}$$

where $\overline{\Omega}_1 = (\frac{1}{\nu}\underline{\Omega}_1^{-1} + B_1'Z_1'Z_1B_1)^{-1}$ and $\overline{\mu}_1 = \left[\frac{1}{\nu}\underline{\Omega}_1^{-1}\underline{\mu}_1 + \left(Z_0 - Z_2B_2A_2' - 2Re(Z_3\bar{B}_\star A_\star') - Z_4\Gamma\right)'Z_1B_1\right]\overline{\Omega}_1$,

- the matrix normal distribution for A_2:

$$p(A_2|\cdot, y) = f_{MN}\left(\overline{\mu}_2, \overline{\Omega}_2, \Sigma\right), \tag{A.5}$$

where $\overline{\Omega}_2 = (\frac{1}{\nu}\underline{\Omega}_2^{-1} + B_2'Z_2'Z_2B_2)^{-1}$ and $\overline{\mu}_2 = \left[\frac{1}{\nu}\underline{\Omega}_2^{-1}\underline{\mu}_2 + \left(Z_0 - Z_1B_1A_1' - 2Re(Z_3\bar{B}_\star A_\star') - Z_4\Gamma\right)'Z_2B_2\right]\overline{\Omega}_2$,

- the matrix normal distribution for $A_{RI} = \left(A_R'A_I'\right)$:

$$p(A_{RI}|\cdot, y) = f_{MN}\left(\overline{\mu}_{RI}, \Sigma, \overline{\Omega}_{RI}\right), \tag{A.6}$$

where $\overline{\Omega}_{RI} = (\frac{2}{\nu}I_{2r_3} + X_{A_{RI}}'X_{A_{RI}})^{-1}$, $\overline{\mu}_{RI} = \overline{\Omega}_{RI}\left[\frac{2}{\nu}\underline{\mu}_{RI} + 2X_{A_{RI}}'\left(Z_0 - Z_1B_1A_1' - Z_2B_2A_2' - Z_4\Gamma\right)\right]$, $\underline{\mu}_{RI} = \left(\underline{\mu}_{\star R}'\underline{\mu}_{\star I}'\right)$ and $X_{A_{RI}} = 2\left[Re(Z_3\bar{B}) - Im(Z_3\bar{B})\right] = 2[(Z_{31} - Z_{32})B_R - (Z_{31} + Z_{32})B_I]$,

- the normal distribution for vector $b_1 = vec(B_1)$:

$$p(b_1|\cdot, y) = f_N\left(\overline{\mu}_{b1}, \overline{\Omega}_{b1}\right), \tag{A.7}$$

where $\overline{\Omega}_{b1} = \left[(m_1 I_{r_1} \otimes P_1^{-1}) + (A_1' \Sigma^{-1} A_1 \otimes Z_1' Z_1)\right]^{-1}$ and $\overline{\mu}_{b1} = \overline{\Omega}_{b1} vec\left[Z_1'\left(Z_0 - Z_2 B_2 A_2' - 2Re(Z_3 \bar{B}_\star A_\star') - Z_4\Gamma\right)\Sigma^{-1} A_1\right]$,

- the normal distribution for vector $b_2 = vec(B_2)$:

$$p(b_2|\cdot, y) = f_N\left(\overline{\mu}_{b2}, \overline{\Omega}_{b2}\right), \tag{A.8}$$

where $\overline{\Omega}_{b2} = \left[(m_2 I_{r_2} \otimes P_2^{-1}) + (A_2' \Sigma^{-1} A_2 \otimes Z_2' Z_2)\right]^{-1}$ and $\overline{\mu}_{b2} = \overline{\Omega}_{b2} vec\left[Z_2'\left(Z_0 - Z_1 B_1 A_1' - 2Re(Z_3 \bar{B}_\star A_\star') - Z_4\Gamma\right)\Sigma^{-1} A_2\right]$,

- the normal distribution for vector $b_R = vec(B_R)$:

$$p(b_R|\cdot, y) = f_N\left(\overline{\mu}_{bR}, \overline{\Omega}_{bR}\right), \tag{A.9}$$

where $\overline{\Omega}_{bR} = \left[x_{bR}'(\Sigma^{-1} \otimes I_T) x_{bR} + 2(m_3 I_{r3} \otimes P_{\star R}^{-1})\right]^{-1}$, $\overline{\mu}_{bR} = 2\overline{\Omega}_{bR} vec\left[Z_{31}' Y_{bR} \Sigma^{-1} A_I - Z_{32}' Y_{bR} \Sigma^{-1} A_R\right]$, $x_{bR} = 2(A_I \otimes Z_{31}) - 2(A_R \otimes Z_{32})$ and $Y_{bR} = Z_0 - Z_1 B_1 A_1' - Z_2 B_2 A_2' - Z_4\Gamma + 2Z_{31} B_I A_R' + 2Z_{32} B_I A_I'$,

- the normal distribution for vector $b_I = vec(B_I)$:

$$p(b_I|\cdot, y) = f_N\left(\overline{\mu}_{bI}, \overline{\Omega}_{bI}\right), \tag{A.10}$$

where $\overline{\Omega}_{bI} = \left[x_{bI}'(\Sigma^{-1} \otimes I_T) x_{bI} + 2(m_3 I_{r3} \otimes (P_{\star R} + P_{\star I} P_{\star R}^{-1} P_{\star I})^{-1})\right]^{-1}$, $\overline{\mu}_{bI} = 2\overline{\Omega}_{bI}$ $vec\left[m_3 (P_{\star R} + P_{\star I} P_{\star R}^{-1} P_{\star I})^{-1} P_{\star I} P_{\star R}^{-1} B_R - Z_{31}' Y_{bI} \Sigma^{-1} A_R + -Z_{32}' Y_{bI} \Sigma^{-1} A_I\right]$, $x_{bI} = -2(A_R \otimes Z_{31}) - 2(A_I \otimes Z_{32})$ and $Y_{bI} = Z_0 - Z_1 B_1 A_1' - Z_2 B_2 A_2' - Z_4\Gamma - 2Z_{31} B_R A_I' + 2Z_{32} B_R A_R'$.

Additionally, if S and ν are estimated, then we obtain:

- the Wishart distribution for S:

$$p(S|\cdot, y) = p(S|\Sigma) = W\left(\left(A_S^{-1} + \Sigma^{-1}\right)^{-1}, q + q_S\right), \tag{A.11}$$

- the inverse gamma distribution for ν:

$$p(\nu|\cdot, y) = p(\nu|\cdot) = iG\left(\bar{s}_\nu, \overline{\Omega}_{bl}\right), \tag{A.12}$$

where $\bar{n}_\nu = \underline{n}_\nu + \frac{n}{2}[n(k-4) + l + r_1 + r_2 + 2r_3]$ and

$$\bar{s}_\nu = \underline{s}_\nu + \frac{1}{2} tr\Big\{\Sigma^{-1}\Big[(\Gamma - \underline{\mu}_\Gamma)' \underline{\Omega}_\Gamma^{-1} (\Gamma - \underline{\mu}_\Gamma) + 2(A_\star - \underline{\mu}_\star)(A_\star - \underline{\mu}_\star)' + + (A_1 - \underline{\mu}_1)\underline{\Omega}_1^{-1}(A_1 - \underline{\mu}_1)' + (A_2 - \underline{\mu}_2)\underline{\Omega}_2^{-1}(A_2 - \underline{\mu}_2)'\Big]\Big\}.$$

Having the set of full conditional posterior distributions, the pseudo-random sample from the joint posterior distribution may be obtained by means of the Gibbs sampler, similarly as Koop et al. (2009) in the CI(1,1) case. In the first step, the initial values are proposed:

$\Sigma^{(0)}$, $\nu^{(0)}$, $\Gamma^{(0)}$, $A_1^{(0)}$, $B_1^{(0)}$, $A_2^{(0)}$, $B_2^{(0)}$, $A_\star^{(0)}$, $B_\star^{(0)}$. Then, the following steps are reiterated (for $s = 1, 2, \ldots$):

- draw $\Sigma^{(s)}$ from the inverse Wishart distribution (A.2),
- draw $\Gamma^{(s)}$ from the matrix normal distribution (A.3),
- draw $A_1^{(s)}$ from the matrix normal distribution (A.4),
- draw $vec(B_1)^{(s)}$ from the normal distribution (A.7) and reshape it to obtain $B_1^{(s)}$,
- calculate $\beta_1^{(s)} = B_1^{(s)}(B_1^{(s)'} B_1^{(s)})^{-\frac{1}{2}}$ and $\alpha_1^{(s)} = A_1^{(s)}(B_1^{(s)'} B_1^{(s)})^{\frac{1}{2}}$,
- draw $A_2^{(s)}$ from the matrix normal distribution (A.5),
- draw $vec(B_2)^{(s)}$ from the normal distribution (A.8) and reshape it to obtain $B_2^{(s)}$,
- calculate $\beta_2^{(s)}$ and $\alpha_2^{(s)}$ as $\beta_2^{(s)} = B_2^{(s)}(B_2^{(s)'} B_2^{(s)})^{-\frac{1}{2}}$ and $\alpha_2^{(s)} = A_2^{(s)}(B_2^{(s)'} B_2^{(s)})^{\frac{1}{2}}$,
- draw $A_R^{(s)}$ and $A_I^{(s)}$ from the matrix normal distribution (A.6),
- draw $vec(B_R)^{(s)}$ from the normal distribution (A.9) and reshape it to obtain $B_R^{(s)}$,
- draw $vec(B_I)^{(s)}$ from the normal distribution (A.10) and reshape it to obtain $B_I^{(s)}$,
- set $A_\star^{(s)} = A_R^{(s)} + iA_I^{(s)}$ and $B_\star^{(s)} = B_R^{(s)} + iB_I^{(s)}$,
- calculate $\beta_\star^{(s)} = B_\star^{(s)}(\bar{B}_\star^{(s)'} B_\star^{(s)})^{-\frac{1}{2}}$ and $\alpha_\star^{(s)} = A_\star^{(s)}(\bar{B}_\star^{(s)'} B_\star^{(s)})^{\frac{1}{2}}$ (note that the square root of the complex Hermitian matrix $(\bar{B}_\star^{(s)'} B_\star^{(s)})^{\frac{1}{2}}$ may be obtained with the Newton method proposed by Higham, 1986),
- check the non-explosive condition: if satisfied, then keep the draws and increase the iteration counter; otherwise, rerun the scheme (note that no new draw for Σ is required).

Additionally, if S and ν are estimated, then:

- draw $S^{(s)}$ from the Wishart distribution given in (A.11),
- draw $\nu^{(s)}$ from the inverse gamma distribution (A.12).

References

Abeysinghe, T., 1994. Deterministic seasonal models and spurious regressions. Journal of Econometrics 61 (2), 259–272.
Ahn, S.K., Cho, S., Chan Seong, B., 2004. Inference of seasonal cointegration: Gaussian reduced rank estimation and tests for various types of cointegration. Oxford Bulletin of Economics and Statistics 66 (2), 261–284.
Andersen, H.H., Højbjerre, M., Sørensen, D., Eriksen, P.S., 1995. The multivariate complex normal distribution. In: Linear and Graphical Models: for the Multivariate Complex Normal Distribution. Springer New York, NY, pp. 15–37.
Arteche, J., 2007. The analysis of seasonal long memory: The case of Spanish inflation. Oxford Bulletin of Economics and Statistics 69 (6), 749–772.
Chern, S.-S., Wolfson, J.G., 1987. Harmonic maps of the two-sphere into a complex Grassmann manifold II. Annals of Mathematics 125 (2), 301–335.
Chikuse, Y., 1990. The matrix angular central Gaussian distribution. Journal of Multivariate Analysis 33 (2), 265–274.
Chikuse, Y., 2003. Statistics on special manifolds, Vol. 1. Springer New York, NY.
Cubadda, G., 1999. Common cycles in seasonal non-stationary time series. Journal of Applied Econometrics 14 (3), 273–291.
Cubadda, G., 2001. Complex reduced rank models for seasonally cointegrated time series. Oxford Bulletin of Economics and Statistics 63 (4), 497–511.
Cubadda, G., Omtzigt, P., 2005. Small-sample improvements in the statistical analysis of seasonally cointegrated systems. Computational Statistics & Data Analysis 49 (2), 333–348.
Díaz-García, J.A., 2013. Distribution theory of quadratic forms for matrix multivariate elliptical distribution. Journal of Statistical Planning and Inference 143 (8), 1330–1342.
Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: Representation, estimation, and testing. Econometrica 55 (2), 251–276.
Engle, R.F., Granger, C.W.J., Hylleberg, S., 1993. Seasonal cointegration: The Japanese consumption function. Journal of Econometrics 55, 275–298.
Franses, P.H., Kunst, R.M., 1999. On the role of seasonal intercepts in seasonal cointegration. Oxford Bulletin of Economics and Statistics 61 (3), 409–433.
Geisel, M.S., 1973. Bayesian comparisons of simple macroeconomic models. Journal of Money, Credit and Banking 5 (3), 751–772.
Gerlach, S., Svensson, L.E., 2003. Money and inflation in the euro area: A case for monetary indicators? Journal of Monetary Economics 50 (8), 1649–1672.
Geweke, J., 2001. Bayesian econometrics and forecasting. Journal of Econometrics 100 (1), 11–15.
Geweke, J., 2005. Contemporary Bayesian econometrics and statistics. John Wiley & Sonss, Inc., Hoboken, New Jersey.
Geweke, J., Amisano, G., 2010. Comparing and evaluating Bayesian predictive distributions of asset returns. International Journal of Forecasting 26 (2), 216–230.
Geweke, J., Amisano, G., 2011. Hierarchical Markov normal mixture models with applications to financial asset returns. Journal of Applied Econometrics 26 (1), 1–29.
Granger, C.W., 1981. Some properties of time series data and their use in econometric model specification. Journal of Econometrics 16 (1), 121–130.
Granger, C.W.J., Siklos, P.L., 1995. Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence. Journal of Econometrics 66 (1-2), 357–369.
Hallman, J.J., Porter, R.D., Small, D.H., 1991. Is the price level tied to the M2 monetary aggregate in the long run? The American Economic Review 81 (4), 841–858.
Hamilton, J.D., 2018. Why you should never use the Hodrick-Prescott filter. Review of Economics and Statistics 100 (5), 831–843.
Hecq, A., 1998. Does seasonal adjustment induce common cycles? Economics Letters 59 (3), 289–297.
Herwartz, H., Reimers, H.-E., 2006. Long-run links among money, prices and output: Worldwide evidence. German Economic Review 7 (1), 65–86.
Higham, N.J., 1986. Newton's method for the matrix square root. Mathematics of Computation 46 (174), 537–549.
Hylleberg, S., Engle, R.F., Granger, C.W., Yoo, B.S., 1990. Seasonal integration and cointegration. Journal of Econometrics 44 (1-2), 215–238.
James, A.T., 1954. Normal multivariate analysis and the orthogonal group. The Annals of Mathematical Statistics 25 (1), 40–75.
Johansen, S., 1995. Likelihood-based inference in cointegrated vector autoregressive models. Oxford University Press.
Johansen, S., Schaumburg, E., 1999. Likelihood analysis of seasonal cointegration. Journal of Econometrics 88 (2), 301–339.
Juselius, K., 2006. The cointegrated VAR model: Methodology and applications. Oxford University Press.
Koop, G., Leon-Gonzalez, R., Strachan, R.W., 2011. Bayesian inference in a time varying cointegration model. Journal of Econometrics 165 (2), 210–220.
Koop, G., León-González, R., Strachan, R.W., 2009. Efficient posterior simulation for cointegrated models with priors on the cointegration space. Econometric Reviews 29 (2), 224–242.
Koop, G., Strachan, R., Van Dijk, H., Villani, M., 2006. Bayesian approaches to cointegration. In: The Palgrave Handbook of Theoretical Econometrics. Palgrave Macmillan, pp. 871–898.
Kotłowski, J., 2005. Money and prices in the Polish economy. Seasonal cointegration approach. Working Papers Series Warsaw School of Economics 3 (05).
Kunst, R.M., 1993. Seasonal cointegration in macroeconomic systems: Case studies for small and large European countries. The Review of Economics and Statistics 75 (2), 325–330.
Kunst, R.M., Franses, P.H., 1998. The impact of seasonal constants on forecasting seasonally cointegrated time series. Journal of Forecasting 17 (2), 109–124.
Lee, H.S., 1992. Maximum likelihood inference on cointegration and seasonal cointegration. Journal of Econometrics 54 (1), 1–47.
Löf, M., Franses, P.H., 2001. On forecasting cointegrated seasonal time series. International Journal of Forecasting 17 (4), 607–621.
Lütkepohl, H., 2005. New introduction to multiple time series analysis. Springer Berlin, Heidelberg.
Meyer, M., Winker, P., 2005. Using HP filtered data for econometric analysis: Some evidence from Monte Carlo simulations. Allgemeines Statistisches Archiv 89 (3), 303–320.
Nerlove, M., 1964. Spectral analysis of seasonal adjustment procedures. Econometrica 32 (3), 241–286.
Pajor, A., Wróblewska, J., 2022. Forecasting performance of Bayesian VEC-MSF models for financial data in the presence of long-run relationships. Eurasian Economic Review 12 (3), 427–428.
Reimers, H.-E., 1997. Forecasting of seasonal cointegrated processes. International Journal of Forecasting 13 (3), 369–380.
Rudebusch, G.D., Svensson, L.E., 2002. Eurosystem monetary targeting: Lessons from U.S. data. European Economic Review 46 (3), 417–442.
Srivastava, A., 2000. A Bayesian approach to geometric subspace estimation. IEEE Transactions on Signal Processing 48 (5), 1390–1400.
Strachan, R.W., van Dijk, H.K., 2007. Bayesian model averaging in vector autoregressive processes with an investigation of stability of the US great ratios and risk of a liquidity trap in the USA, UK and Japan. Report / Econometric Institute, Erasmus University Rotterdam (No. EI 2007-11).
Tödter, K.-H., Reimers, H.-E., 1994. P-star as a link between money and prices in Germany. Weltwirtschaftliches Archiv 130 (2), 273–289.
Villani, M., 2006. Bayesian point estimation of the cointegration space. Journal of Econometrics 134 (2), 645–664.
Wesche, K., 1997. The stability of European money demand: An investigation of M3H. Open Economies Review 8 (4), 371–391.
Wooding, R.A., 1956. The multivariate distribution of complex normal variables. Biometrika 43 (1/2), 212–215.
Wróblewska, J., 2009. Bayesian model selection in the analysis of cointegration. Central European Journal of Economic Modelling and Econometrics 1 (1), 57–69.
Wróblewska, J., 2020. A note on some extensions of the matrix angular central gaussian distribution. Preprint on arXiv at arXiv:2010.03243.