Appendix

Appendix A: Estimation of a VARMA Model

A K dimensional VARMA(p,q) process can be written as

$$\mathbf{X}_t = \mathbf{c} + \mathbf{\Phi}_1\mathbf{X}_{t-1} + \ldots + \mathbf{\Phi}_p\mathbf{X}_{t-p} + \boldsymbol{\upsilon}_t - \mathbf{\Theta}_1\boldsymbol{\upsilon}_{t-1} - \ldots - \mathbf{\Theta}_q\boldsymbol{\upsilon}_{t-q}, \tag{11}$$

To represent the K-dimensional VARMA(p,q) process of (11) in terms of K-SCMs, K linear transformation are preformed via the transformation matrix resulting in

$$\mathbf{z}_t = \mathbf{B}\mathbf{X}_t \tag{12}$$

where $\mathbf{B} = (\beta_1, \beta_2, \ldots, \beta_k)'$ is a $(K \times K)$ invertible matrix while $\mathbf{z}_t = (z_{1,t}, z_{2,t} \ldots, z_{K,t})'$ is a transformed process associated with K-SCM(p_i, q_i) for $i = 1, 2, \ldots, K$. A non-zero linear combination $z_{i,t} = \beta_i' X_{i,t}$, follows a SCM$(p_i, q_i)$ if β_i has the following properties:

$$\begin{aligned}
\beta_i'\mathbf{\Phi}_{p_i} &\neq \mathbf{0}^T \text{ where } 0 \leq p_i \leq p,\\
\beta_i'\mathbf{\Phi}_l &= \mathbf{0}^T \text{ for } l = p_i + 1, \ldots, p,\\
\beta_i'\mathbf{\Theta}_{q_i} &\neq \mathbf{0}^T \text{ where } 0 \leq q_i \leq q,\\
\beta_i'\mathbf{\Theta}_l &= \mathbf{0}^T \text{ for } l = q_i + 1, \ldots, q.
\end{aligned}$$

The scalar random variable $z_{i,t}$, depends only on lags 1 to p_i of all variables and lags 1 to q_i of all innovations in the system. The aim of identifying scalar components is to examine whether any simplifying embedded structures underlying this process can provide a parsimonious VARMA structure.

Embedded scalar component models are determined through a series of canonical correlation tests. Let the estimated squared canonical correlations between $\mathbf{Y}_{m,t} \equiv \left(\mathbf{y}_t', \ldots, \mathbf{y}_{t-m}'\right)$ and $\mathbf{Y}_{h,t-1-j} \equiv \left(\mathbf{y}_{t-1-j}', \ldots, \mathbf{y}_{t-1-j-h}'\right)'$ be $\widehat{\lambda}_1 < \widehat{\lambda}_2 < \ldots < \widehat{\lambda}_K$. As suggested by Tiao & Tsay (1989), the test statistic for at least s SCM(p_i, q_i), i.e., s insignificant canonical correlations, against the alternative of less than s scalar components is

$$C(s) = -(n-h-j)\sum_{i=1}^{s} \ln\left\{1 - \frac{\widehat{\lambda}_i}{d_i}\right\} \overset{a}{\sim} \chi^2_{s\times\{(h-m)K+s\}} \tag{13}$$

where d_i is a correction factor that accounts for the fact that the canonical variates could be moving averages of order j and it is calculated as follows:

$$d_i = 1 + 2\sum_{v=1}^{j} \widehat{\rho}_v\left(\widehat{\mathbf{r}}_i'\mathbf{Y}_{m,t}\right)\widehat{\rho}_v\left(\widehat{\mathbf{g}}_i'\mathbf{Y}_{h,t-1-j}\right) \tag{14}$$

where $\widehat{\rho}_v(.)$ is the v^{th} order autocorrelation of its argument and $\widehat{\mathbf{r}}_i'\mathbf{Y}_{m,t}$ and $\widehat{\mathbf{g}}_i'\mathbf{Y}_{h,t-1-j}$ are the canonical variates corresponding to the i^{th} canonical correlation between $\mathbf{Y}_{m,t}$ and $\mathbf{Y}_{h,t-1-j}$. Let, $\mathbf{\Gamma}(m,h,j) = E(\mathbf{Y}_{h,t-1-j}\mathbf{Y}_{m,t}^{'})$. This is a sub-matrix of the Hankel matrix of the autocovariance matrices of $\mathbf{y}_t$. Note that zero canonical correlations imply and are implied by $\mathbf{\Gamma}(m,h,j)$ having a zero eigenvalue.

The modelling of the VARMA(p,q) process is carried out in three stages, and they are described as follows.[1]

Stage I: Identification of the SCMs

By strategically choosing $\mathbf{Y}_{m,t}$ and $\mathbf{Y}_{h,t-1-j}$, the overall tentative order of the VARMA(p,q) is identified. The process begins by searching for K SCMs of the most parsimonious possibility, i.e., SCM$(0,0)$, which is a white noise process by testing for the rank of $\mathbf{\Gamma}(0,0,0) = E(\mathbf{Y}_{0,t-1}\mathbf{Y}_{0,t}^{'})$; where $\mathbf{Y}_{m,t} = \mathbf{Y}_{0,t}$ and $\mathbf{Y}_{h,t-1-j} = \mathbf{Y}_{0,t-1}$. If we do not find K linearly

[1] For further details, refer to Athanasopoulos & Vahid (2008b) and Tiao & Tsay (1989).

independent white noise scalar processes, we set $m = h$, and by incrementing m and j, we search for the next set of K linearly independent scalar components. Once the overall tentative order (p, q) is identified, we repeat the search process to identify the individual components. Starting again from the most parsimonious SCM(0,0), we sequentially search for K linearly independent vectors $(\alpha_1, \ldots, \alpha_K)$ for $m = 0, \ldots, p$, $j = 0, \ldots, q$ and $h = m + (q - j)$.

The test results from identifying the overall tentative order and the individual SCMs are tabulated in Table A.1, referred to as *Criterion* and *Root* tables. We report the results of all canonical correlations test statistics divided by their χ^2 critical for the full period. If the entry in the $(m, j)^{th}$ cell is less than one in the "Criterion Table", it shows that there are six SCMs of order (m, j) or lower in this system. From Panel A, the identified overall order of the system is VARMA$(2, 1)$.

Table A.1: Stage I of the identification process of a VARMA model with DivM3

PANEL A: Criterion Table

m \ j	0	1	2	3	4
0	271.33[a]	25.87	13.91	9.41	7.05
1	11.94	3.20	1.80	1.16	0.66
2	2.43	**0.87**	0.81	0.82	0.73
3	0.88	0.92	0.92	0.84	0.87
4	1.14	1.06	0.97	0.98	0.97

PANEL B: Root Table

m \ j	0	1	2	3	4
0	0	0	0	0	1
1	0	**3**	4	5	6
2	4	**6**	9	10	11
3	6	10	12	15	16
4	5	11	16	18	21

[a] *The statistics are normalized by the corresponding 5% χ^2 critical values*

Conditional on the overall order of $(2, 1)$, canonical correlation tests are performed to identify the individual orders of embedded SCMs. The number of insignificant canonical correlations found are tabulated in Panel B of Table A.1. This is referred to as the "Root Table"and the figures in bold show that three SCMs of order $(1, 1)$ is identified in position $(m, j) = (1, 1)$. Then, there are six SCMs of order $(2, 1)$ at position $(m, j) = (2, 1)$. For every SCM(m, j) nest all scalar components of order $(\leq m, \leq j)$ and for each individual SCM$(p_1 < p, q_1 < q)$, there will be $\xi = \min\{m - p_1 + 1, j - q_1 + 1\}$ zero canonical correlations at position $(m \geq p_1, j \geq q_1)$. Hence, a new SCM(m, j) is found for every increment above ξ. For the identified six SCMs of order $(2, 1)$, only three are new, while the other three are carried over from the SCM$(1, 1)$. So, our identified VARMA$(2, 1)$ consists of three SCM$(1, 1)$ and three SCM$(2, 1)$.

If the identified K linearly independent scalar components are characterized by the transformation matrix $\mathbf{B} = (\beta_1, \beta_2, \ldots, \beta_k)'$, the system in (11) can be rotated to obtain

$$\mathbf{z}_t - \mathbf{\Phi}_1^* \mathbf{z}_{t-1} - \ldots - \mathbf{\Phi}_p^* \mathbf{z}_{t-p} = \mathbf{u}_t - \mathbf{\Theta}_1^* \mathbf{u}_{t-1} - \ldots - \mathbf{\Theta}_q^* \mathbf{u}_{t-q}, \tag{15}$$

where $\mathbf{u}_t = \mathbf{B}\upsilon_t$, $\mathbf{\Phi}_j^* = \mathbf{B}\mathbf{\Phi}_j\mathbf{B}^{-1}$ and $\mathbf{\Theta}_j^* = \mathbf{B}\mathbf{\Theta}_j\mathbf{B}^{-1}$ for $j = 1$ to p(or q).

In the rotated model, each row represents one identified SCM(p_i, q_i). However, obtaining the orders of SCMs does not necessarily lead to a uniquely identified system. For example, if two scalar components were identified such that $z_{r,t} = SCM\,(p_r, q_r)$ and $z_{s,t} = SCM\,(p_s, q_s)$, where $p_r > p_s$ and $q_r > q_s$, the system will not be identified as we need to set $\min\{p_r - p_s, q_r - q_s\}$ autoregressive or moving average parameters to zero. This process is known as the "general rule of elimination," and in order to identify a canonical VARMA model, we set the moving average parameters to zero.

Stage II: Identification of the transformation matrix B

The space spanned by $\mathbf{z}_{t-1}$ to $\mathbf{z}_{t-p}$ is the same as the space spanned by $\mathbf{X}_{t-1}$ to $\mathbf{X}_{t-p}$. So, for the transformed model (15), the right hand side of the equation can be written in terms of $\mathbf{X}_{t-1}$ to $\mathbf{X}_{t-p}$ instead of $\mathbf{z}_{t-1}$ to $\mathbf{z}_{t-p}$ without affecting the restrictions imposed by the scalar component rules.[2] Hence, if we rotate the system by replacing $\mathbf{z}_{t-1}, \ldots, \mathbf{z}_{t-p}$ with $\mathbf{BX}_{t-1}, \ldots, \mathbf{BX}_{t-p}$, the system can be represented in terms of the original series as follows:

$$\mathbf{BX}_t = \mathbf{\Gamma}_1\mathbf{X}_{t-1} + \ldots + \mathbf{\Gamma}_p\mathbf{X}_{t-p} + \mathbf{u}_t - \mathbf{\Theta}_1^*\mathbf{u}_{t-1} - \ldots - \mathbf{\Theta}_q^*\mathbf{u}_{t-q}, \tag{16}$$

where $\mathbf{\Gamma}_i = \mathbf{\Phi}_1^*\mathbf{B}$ for $i = 1, \ldots, p$ and with $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p$ and $\mathbf{\Phi}_1^*, \ldots, \mathbf{\Phi}_p^*$ satisfying the same restrictions as the right hand side of equation (15).

Some of the parameters in $\mathbf{B}$ are redundant and can be eliminated. A brief description about the rules of placing restrictions on the redundant parameters are as follows:

1. Each row of the transformation matrix $\mathbf{B}$ can be multiplied by a constant without changing the structure of the model; i.e, one parameter in each row can be normalized to one as long as this parameter is not zero. To make sure of this tests of predictability using subsets of variables are performed.

2. Any linear combination of a SCM(p_1, q_1) and a SCM(p_2, q_2) is a SCM$(\max\{p_1, p_2\}, \max\{q_1, q_2\})$. For all cases where there are two SCMs with weakly nested orders, i.e., $p_1 \geq p_2$ and $q_1 \geq q_2$, if the parameter in the i^{th} column of the row of $\mathbf{B}$ corresponding to the SCM(p_2, q_2) is normalized to one, the parameter in the same position in the row of $\mathbf{B}$ corresponding to SCM(p_1, q_1) should be restricted to zero.

Detailed explanations on these issues, together with examples, can be found in Athanasopoulos & Vahid (2008b).

Stage III: Estimation of the uniquely identified system
The identified model is estimated using FIML and is given by

$$\ln \text{Ł}(\mathbf{A}, \mathbf{\Sigma}) = -\frac{N-p}{2}(-\ln|\mathbf{B}| + \ln|\mathbf{\Sigma}_\varepsilon| - \ln|\mathbf{B}'|) - \frac{1}{2}\varepsilon_t'\mathbf{\Sigma}_\varepsilon^{-1}\varepsilon_t \tag{17}$$

thus

$$\ln \text{Ł}(\mathbf{A}, \mathbf{\Sigma}) \propto (N-p)\ln|\mathbf{B}| - \frac{N-p}{2}\ln|\mathbf{\Sigma}_\varepsilon| - \frac{1}{2}\varepsilon_t'\mathbf{\Sigma}_\varepsilon^{-1}\varepsilon_t \tag{18}$$

where $\mathbf{A} = [\mathbf{B} : \mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p : \mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_q]$ and $\mathbf{\Sigma} = var(\mathbf{X}_t/\mathbf{X}_{t-1}, \ldots, \mathbf{X}_1)$. As in Hannan & Rissanen (1982), a long VAR is used to obtain initial values of the parameters.

[2] A detailed explanation on this can be found in Athanasopoulos & Vahid (2008b).

Appendix B: US Business Cycle Variables

Table B.1: Data Descriptions and Sources

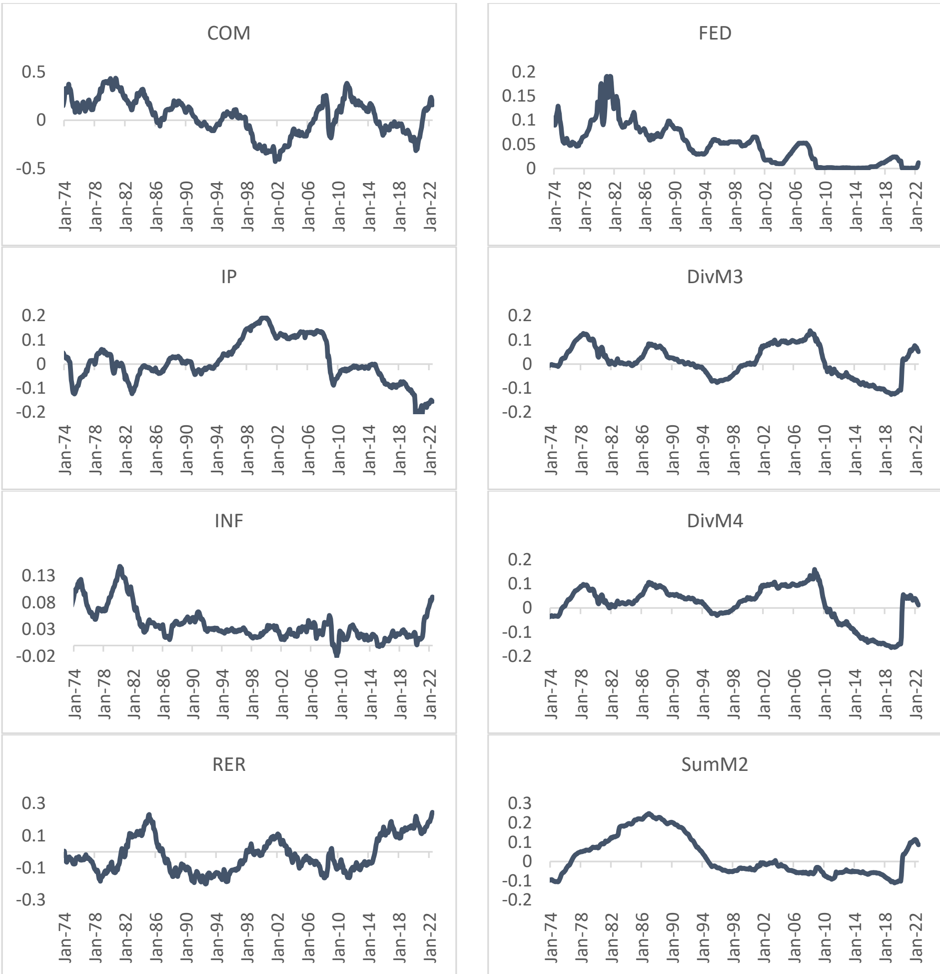
Variable	Description	Source
COM_t	Commodity Price Index	CRB/BLS spot index
	(SA, logs and detrended)	(Commodity Research Bureau)
IP_t	US Industrial Production Index	FRED Database
	(SA, logs and detrended)	
INF_t	US Consumer Price Index	FRED Database
	(SA, logs and % change per annum)	
$SumM2_t$	Simple Sum of M2	FRED Database
	(SA, logs and detrended)	
$DivM3_t$	Divisia M3	Centre for Financial Stability
	(SA, logs and detrended)	
$DivM4_t$	Divisia M4	Centre for Financial Stability
	(SA, logs and detrended)	
FED_t	Federal Funds Rate (%)	FRED Database
RER_t	Real Narrow Effective Exchange rate index	FRED Database
	(SA, logs and detrended)	

Table B.2: Unit Root Test Results

Series	ADF		PP	
COM_t	-2.55	(0.305)	-2.29	(0.438)
IP_t	-1.27	(0.894)	-1.48	(0.833)
INF_t	-2.69	(0.077)	-2.36	(0.153)
$SumM2_t$	-2.74	(0.217)	-1.88	(0.633)
$DivM3_t$	-2.81	(0.194)	-1.87	(0.868)
$DivM4_t$	-2.74	(0.222)	-2.18	(0.498)
FED_t	-2.08	(0.252)	-1.89	(0.336)
RER_t	-1.73	(0.737)	-1.75	(0.724)

Note(s): ADF is the Augmented Dickey-Fuller test, with augmentation selected by AIC. The Phillips-Perron test is applied with the Bartlett kernel and automatic Newey-West bandwidth selection. Values in parentheses are P-values.

Figure B.1: Data Series



Note: Notation is defined as in Table B.1. The data series are discussed in Subsection 3.1.

References

Athanasopoulos, G., & Vahid, F. (2008b). A complete VARMA modelling methodology based on scalar components. *Journal of Time Series Analysis*, *29*, 533–554.

Hannan, E., & Rissanen, J. (1982). Recursive estimation of autoregressive-moving average order. *Biometrica*, *69*, 81–94.

Tiao, G. C., & Tsay, R. S. (1989). Model specification in multivariate time series. *Journal of Royal Statistical Society, Series B (Methodological)*, *51(2)*, 157–213.