The Vector Error Correction Index Model: Representation and Statistical Inference

Gianluca Cubadda Marco Mazzali

Università di Roma "Tor Vergata"

June 24, 2022

Introduction

- Since the seminal paper of Sims (1980), Vector Autoregressive models (VARs) have been largely used for both forecasting and structural analysis. However, the number of parameters to be estimated increases with the square of the number of variables in the system. This issue is known as the *curse of dimensionality*.
- The most successful alternatives in empirical applications are the Dynamic Factor Models (DFMs), Bayesian VARs (BVARs), and, lately, sparse VARs.
- An attractive feature of DFMs is that few latent factors drive the entire economy, see *e.g.* Lippi (2018, 2019), Stock and Watson (2016) and the references therein.
- Both BVARs and sparse VARs employ parameter shrinkage methods and have proven to be useful in handling large datasets.

Continued...

- In the classical VAR framework, the Likelihood Ratio Test (LRT) (Johansen (1995)) seriously over-rejects the null when the number of series n is not small and the sample size T is not much larger than those typically employed in macroeconomic applications, see *i.a.* Ho and Sorensen (1996) and Gonzalo and Pitarakis (1999).
- When both n and T diverge, Onatski and Wang (2018) proved that the LRT statistic does not converge to the usual asymptotic distribution, whereas Bykhovskaya and Gorin (2022) proposed a variant of the LRT for a VAR(1) model and derived its limit distribution.
- We instead focus on a medium-n framework, given the diffuse evidence that no substantial gains are obtained by increasing the dimension of the VAR beyond 20 (see, *i.a.*, Bańbura et al. (2010), Giannone et al. (2008), Koop (2013), and Cubadda and Guardabascio (2019)). Hence, we don't assume that n diverges for inference.

Continued...

- Recently, there has been a renewed interest in the Multivariate Autoregressive Index Model (MAI) by Reinsel (1983). The MAI is obtained by imposing a particular reduced rank structure on a stationary VAR such that time series are driven by the lags of the *indexes*, i.e. few linear combinations of the variables, which can be considered as "observable" factors.
- The MAI has recently been extended in several directions, see *i.a.*, Carriero et al. (2016, 2022), Cubadda et al. (2017), Cubadda and Guardabascio (2019), and Cubadda and Hecq (2022). The present paper aims at allowing for cointegration (Engle and Granger, 1987) in the MAI framework.

Continued...

- We achieve dimension reduction by proposing a new modeling such that the first differences of a set of cointegrated time series are endowed with an index structure. We label the resulting specification as the Vector Error-Correction Index Model (VECIM).
- An interesting property of the VECIM is that the indexes themselves follow a VECM model with a dimension smaller than n.
- For estimation, we propose a switching algorithm that increases the Gaussian likelihood in each step.
- In line with previous contributions indicating that information criteria (IC) are particularly useful in selection of VAR models (see e.g. Nielsen (2006), Cavaliere et al. (2015), and Cavaliere et al. (2018)), we rely on IC rather than on testing procedures for the specification of the VECIM.

Representation theory

- We assume that series Y_t, whose elements are at most I(1), follow the Vector Error-Correction Model (VECM):

$$\Delta Y_t = \alpha_0 \beta' Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \varepsilon_t \tag{1}$$

  where α_0 and β are full-rank $n \times r$ $(r < n)$ matrices, Γ_j is an $n \times n$ matrix for $j = 1, ..., p-1$, ε_t is an n-vector of errors such that $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{E}(\varepsilon_t \varepsilon_t') = \Omega$ is a finite and positive definite matrix, $\mathrm{E}(\varepsilon_t | F_{t-1}) = 0$ and F_t is the natural filtration of the process Y_t.
- Under such assumptions, it is well known that series Y_t are $\mathrm{CI}(1,1)$, see *e.g.* Johansen (1995).

Continued...

- Based on Reinsel (1983), we further assume that the system is endowed with the following index structure

$$\Delta Y_t = \alpha_0 \gamma' \omega' Y_{t-1} + \sum_{j=1}^{p-1} \alpha_j \omega' \Delta Y_{t-j} + \varepsilon_t \tag{2}$$

  where ω is a full-rank $n \times q$ matrix with $q \geq r$, γ is a full-rank $q \times r$ matrix such that $\beta = \omega\gamma$, and α_j is an $n \times q$ matrix for $j = 1, ..., p-1$.
- We label (2) as the Vector Error Correction Index Model (VECIM) and variables $f_t = \omega' Y_t$ as the indexes. The interpretation is that past information influences the first differences ΔY_t through a limited number of channels (q over n).
- The VECIM has $(np + r - n - q)(n - q)$ parameters less than the VECM, which leads to a rather more parsimonious specification when q is small w.r.t. n.

Continued...

- The VECIM has the following implication for the VAR representation of series Y_t

$$\Delta Y_t = \sum_{j=1}^{p} \theta_j \omega' Y_{t-j} + \varepsilon_t,$$

where $\theta_{j+1} = \alpha_{j+1} - \alpha_j$ for $j = 1, ..., p-1$ with $\alpha_p = 0$, and $\theta_1 = \alpha_0 \gamma' - \sum_{j=2}^{p} \theta_j$.

- Moreover, premultiplying both sides of the VECIM with ω' we get

$$\Delta f_t = \underline{\alpha}_0 \gamma' f_{t-1} + \sum_{j=1}^{p-1} \underline{\alpha}_j \Delta f_{t-j} + \epsilon_t \tag{3}$$

where $\underline{\alpha}_0 = \omega' \alpha_0$, $\underline{\alpha}_j = \omega' \alpha_j$ for $j = 1, ..., p-1$, and $\epsilon_t = \omega' \varepsilon_t$. Hence, the indexes f_t follow a VECM model with a cointegration matrix equal to γ.

- When $r = 0$ series ΔY_t are generated by a stationary MAI, whereas when $r = q$ series Y_t follow a VECIM with $\beta = \omega$.

Continued...

- Given the Wold representation

$$\Delta Y_t = \Psi(L)\varepsilon_t$$

and inserting between $\Psi(L)$ and ε_t the decomposition of the identity matrix as in Centoni and Cubadda (2003)

$$\Omega\omega(\omega'\Omega\omega)^{-1}\omega' + \omega_\perp(\omega'_\perp\Omega^{-1}\omega_\perp)^{-1}\omega'_\perp\Omega^{-1} = I_n$$

we obtain the decomposition $Y_t = \chi_t + \iota_t$ where

$$\begin{aligned} \Delta\chi_t &= C(L)\epsilon_t, \\ \Delta\iota_t &= \Psi(L)\omega_\perp(\omega'_\perp\Omega^{-1}\omega_\perp)^{-1}\xi_t, \end{aligned} \tag{4}$$

$C(L) = \Psi(L)\Omega\omega(\omega'\Omega\omega)^{-1}$, and $\xi_t = \omega'_\perp\Omega^{-1}\varepsilon_t$.

- Since $\mathrm{E}(\epsilon'_t\xi_t) = 0$ and χ_t has the same innovations as the indexes f_t, we can interpret the former as the common component of series Y_t. Hence, we can use the methods adopted in the structural FAVAR (Bernanke et al., 2005) literature to recover the structural shocks from the common shocks ϵ_t only.

Continued...

- Moreover, inserting between $C(L)$ and ϵ_t in Equation (4) the following decomposition of the identity matrix

$$\Sigma\underline{\alpha}_{0\perp}(\underline{\alpha}'_{0\perp}\Sigma\underline{\alpha}_{0\perp})^{-1}\underline{\alpha}'_{0\perp} + \underline{\alpha}_0(\underline{\alpha}'_0\Sigma^{-1}\underline{\alpha}_0)^{-1}\underline{\alpha}'_0\Sigma^{-1} = I_q$$

where $\Sigma = \omega'\Omega\omega$, one obtains $\chi_t = \pi_t + \tau_t$, where

$$\Delta\pi_t = \underbrace{C(L)\Sigma\underline{\alpha}_{0\perp}(\underline{\alpha}'_{0\perp}\Sigma\underline{\alpha}_{0\perp})^{-1}}_{P(L)}\underbrace{\underline{\alpha}'_{0\perp}\epsilon_t}_{u_t},$$

$$\Delta\tau_t = \underbrace{C(L)\underline{\alpha}_0(\underline{\alpha}'_0\Sigma^{-1}\underline{\alpha}_0)^{-1}}_{T(L)}\underbrace{\underline{\alpha}'_0\Sigma^{-1}\epsilon_t}_{\eta_t},$$

- Since $\mathrm{E}(u'_t\eta_t) = 0$ and u_t are the innovations of the common trends of the indexes f_t, we see that the common component χ_t is further separated into a common permanent component π_t and a common transitory component τ_t.

Statistical Inference

Based on Boswijk (1995) and Cubadda et al. (2017), we offer some switching algorithms that increase the Gaussian likelihood of the VECIM in each step. When $1 < r < q$, the procedure goes as follows:

1. Given (initial) estimates of both γ and ω, maximize the conditional Gaussian likelihood $\mathcal{L}(\alpha, \Omega | \gamma, \omega)$ by estimating $\alpha = [\alpha_0, \alpha_1, ..., \alpha_{p-1}]$ and Ω with OLS on the VECIM.
2. Premultiply by $\Omega^{-1/2}$ and apply the Vec operator to both the sides of the VECIM to get

$$\begin{aligned} & \Omega^{-1/2}\Delta Y_t - \Omega^{-1/2}\varepsilon_t \qquad (5) \\ = & \left(Y'_{t-1} \otimes \Omega^{-1/2}\alpha_0\gamma' + \sum_{j=1}^{p-1} \Delta Y'_{t-j} \otimes \Omega^{-1/2}\alpha_j \right) \mathrm{Vec}(\omega') \end{aligned}$$

Given the previously obtained estimates of α, γ, and Ω, maximize $\mathcal{L}(\omega | \alpha, \gamma, \Omega)$ by estimating $\mathrm{Vec}(\omega')$ with OLS on Equation (5).

Continued...

3. Given the previously obtained estimates of ω, maximize $\mathcal{L}(\gamma|\omega)$ by estimating γ as the eigenvectors that correspond to the r largest eigenvalues of the matrix

$$S_{11}^{-1} S_{10} S_{00}^{-1} S_{01}$$

where $S_{ij} = \sum_{t=p+1}^{T} R_{i,t} R_{j,t}'$ for $i, j = 0, 1$, $R_{0,t}$ and $R_{1,t}$ are respectively the residuals of an OLS regression of ΔY_t and $\omega' Y_{t-1}$ on $[\Delta Y_{t-1}'\omega, ..., \Delta Y_{t-p+1}'\omega']'$.

4. Repeat steps 1 to 3 till numerical convergence occurs.

- Modifications of the algorithm are necessary when $q = 0, r$.

Continued...

In order to specify the values for p, r and q, two alternative approaches are in principle viable:

1. Sequential LR testing in order to fix initially p, then r conditionally on p and finally q conditionally on p and r.
2. Simultaneously establishing the triple p, r, q by means of IC. In particular, model (2) is estimated for all plausible values of p, r, q. and we select the triple that minimizes an IC such as AIC, HQIC, or BIC.

- Although sequential testing is often used to specify the VECM, we find more convincing the idea of jointly determining p, r, q in one single search.

Monte Carlo analysis

- We simulate a VAR(3) process with $q = 1$ and $r = 0, 1$; $q = 3$ and $r = 0, 1, 3$ for $n = 6, 12, 18$ variables, and with $q = 5$ and $r = 0, 1, 3, 5$ for $n = 12, 18$ variables. The number of observations is $T = 200, 400, 600$.
- The VAR representations of the indexes has q pairs of complex conjugate roots with moduli that are included in $[1.\bar{1}, 1.\bar{6}]$ and angular frequencies that belong to the business cycle band for quarterly data. Moreover, it has $q - r$ roots equal to 1, and r real roots that are included in $[1.\bar{1}, 1.\bar{6}]$. Hence, the DGP is constructed to mimic a system of $I(1)$ quarterly time series with common trends and common business cycles.
- The idiosyncratic component follows instead an $(n - p)-$dimensional random walk.

Continued...

- The VECIM is evaluated by means of three statistics: (i) the percentage of correct estimation of the number of indexes; (ii) the Frobenius Norm between the reduced-form matrix of coefficients obtained with the VECIM and the true ones, relative to the Frobenius Norm obtained with the VECM estimated by the usual ML procedure; (iii) the Mean Square Forecast Error relative to the one-step-ahead forecast of the VECM (RMSFE). Results are based on 500 replications for each combination.
- HQIC performs better than both AIC and BIC in identification. This outcome is robust to variations of n, T, r and q.
- The VECIM, as identified by the HQIC, outperforms the VECM both in terms of estimation precision and prediction.
- The absolute performances of the VECIM are negatively affected by system the dimension, although much less than the VECM.

Empirics I: Contribution of shocks to the US business cycle

We use ten quarterly US time series for the period $1955Q1 - 2020Q1$, which are transformed such that each variable is at most $I(1)$:

- (i) real GDP per capita; (ii) real consumption per capita, computed as the sum of non-durable consumption and services; (iii) real investment per capita, computed as the sum of investment and durable consumption; (iv) hours worked per person; (v) inflation, obtained from the difference of the log of the GDP deflator; (vi) the unemployment rate, (vii) the nominal interest rate, as measured by the FED Funds rate; (viii) TFP, which is the cumulated sum of the utilization-adjusted TFP as computed in Fernald (2014), (ix) the non-farm business sector labor productivity; (x) labor share.
- We take the log of real variables, the first difference of the log of nominal variables, the remaining variables are unmodified.

Continued...

- Given the results of the Monte Carlo, we rely on HQIC to specify the VECIM and set $p = 2$, $r = 3$ and $q = 4$.
- We decompose variables in two components, one of which is common, χ_t (which is further decomposed into a permanent component π_t and a transitory one τ_t) and the other is idiosyncratic, ι_t:

$$Y_t = \chi_t + \iota_t = \pi_t + \tau_t + \iota_t$$

- Furthermore, being all the components above not cross-correlated at any lead and lag, we can rewrite the above decomposition in terms of their (pseudo) spectral densities:

$$F(\lambda) = F_\chi(\lambda) + F_\iota(\lambda) = F_\pi(\lambda) + F_\tau(\lambda) + F_\iota(\lambda), \lambda \in (0, \pi]$$

- Notice that $F(\lambda)$ and its components can be easily obtained from the VECIM parameters.

Continued...

Table: Variance contribution of each component at the business cycle frequency band, i.e. $\lambda \in [2\pi/32, 2\pi/6]$

	Common	Permanent	Transitory
Unemployment	96.8 [94.6, 98.2]	40.8 [27.9, 55.0]	59.2 [45.0, 72.1]
Output	94.8 [91.6, 97.1]	35.4 [23.1, 48.2]	64.6 [51.8, 76.9]
Hours Worked	89.8 [84.6, 93.3]	41.7 [27.3, 56.1]	58.3 [43.9, 72.7]
Investment	94.8 [91.8, 96.7]	35.4 [22.9, 48.3]	64.6 [51.7, 77.1]
Consumption	94.1 [87.7, 97.7]	49.3 [32.0, 66.7]	50.7 [33.3, 68.0]
TFP	57.8 [43.0, 71.3]	61.2 [26.6, 84.5]	38.8 [15.5, 73.4]
Labor Prod.	74.5 [62.8, 83.8]	42.1 [23.1, 59.4]	57.9 [40.6, 76.9]
Labor Share	83.5 [72.5, 91.4]	28.3 [15.2, 44.2]	71.7 [55.8, 84.8]
Inflation	86.2 [78.1, 91.7]	51.7 [22.3, 74.6]	48.3 [25.4, 77.7]
Nom. Int. rate	83.1 [75.1, 88.7]	41.6 [22.1, 63.4]	58.4 [36.6, 77.9]

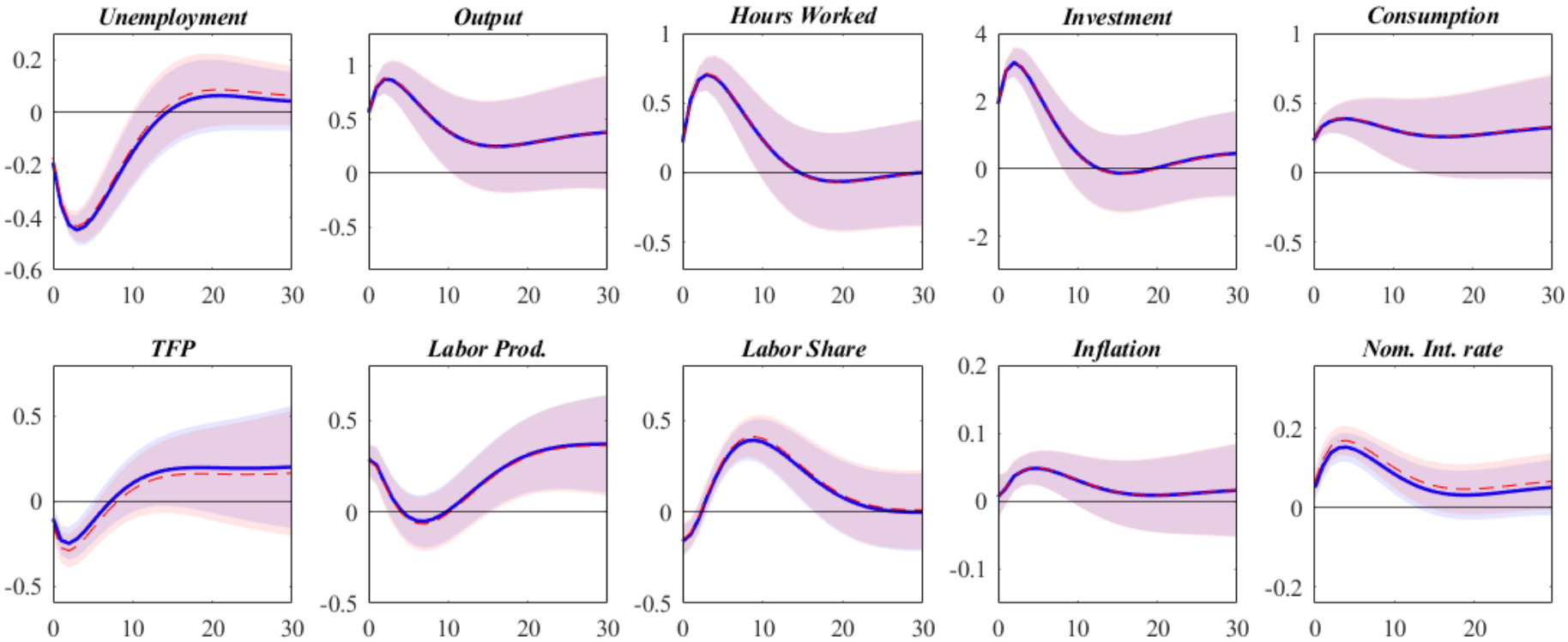
Notes: the table compares the median of the variance contribution of each component and, in brackets, the $16th$ and $84th$ percentiles. The results are obtained with bootstrap.

Empirics II: Identification of the business cycle driver

- Based on Angeletos et al. (2020), we aim at identifying the shock that maximizes the variability of a key pro-cyclical variable at the business cycle frequency. This is labelled as *Main Business Cycle* shock (MBC).
- Similarly, we can apply the same procedure on the common component only and obtain the *Main Common Business Cycle* shock (MCBC). The motivation lies in the fact that the presence of idiosyncratic errors, usually interpreted as measurement errors or sectoral shocks, may contaminate the estimates and consequently lead to misinterpretations, see *i.a.* Forni et al. (2009, 2020), Lippi (2021), and Cubadda and Hecq (2022).
- Following Angeletos et al. (2020), we use the unemployment rate as the key series for which the variability [of its common component] at the business cycle frequencies is maximized by the MBC [MCBC].

Continued...

Figure: Comparison between MBC and MCBC impulse responses



Notes: the red dotted line is the point estimation of the responses to the *main common business cycle shock*, the blue solid line is the point estimation of the responses to the *main business cycle shock* whereas the shaded areas are the 1stde confidence bands obtained with bootstrap

Continued...

Table: Measures of contribution - MBC vs MCBC

	MBC		MCBC	
	Short-Run	Long-Run	Short-Run	Long-Run
Unemployment	68.5	12.0	68.8	17.0
	[58.0, 80.6]	[1.6, 37.8]	[57.8, 81.0]	[2.5, 47.5]
Output	58.3	14.0	62.9	19.3
	[45.2, 69.9]	[1.4, 39.0]	[49.7, 75.0]	[2.1, 48.3]
Hours Worked	59.2	15.0	68.9	22.4
	[42.9, 70.5]	[1.7, 40.1]	[51.7, 80.2]	[3.3, 54.9]
Investment	59.4	16.2	64.0	24.4
	[48.1, 71.8]	[1.8, 41.9]	[51.8, 76.8]	[3.4, 56.6]
Consumption	40.3	13.5	45.2	16.9
	[20.2, 55.1]	[1.7, 36.3]	[25.1, 60.5]	[2.2, 43.6]
TFP	16.4	8.9	38.1	24.8
	[6.7, 28.5]	[0.9, 26.7]	[16.5, 58.1]	[3.6, 54.5]
Labor Prod.	28.3	6.4	40.7	11.1
	[15.5, 41.7]	[0.6, 24.8]	[23.1, 56.1]	[1.1, 40.8]
Labor Share	25.5	10.2	32.4	15.7
	[14.9, 38.1]	[1.1, 33.5]	[18.8, 46.0]	[2.0, 46.5]
Inflation	18.5	6.9	21.3	17.4
	[6.0, 35.6]	[0.7, 24.3]	[6.9, 40.1]	[2.1, 46.9]
Nom. Int. rate	51.6	9.8	68.3	19.4
	[24.7, 67.5]	[1.1, 31.3]	[36.7, 83.2]	[2.3, 51.6]

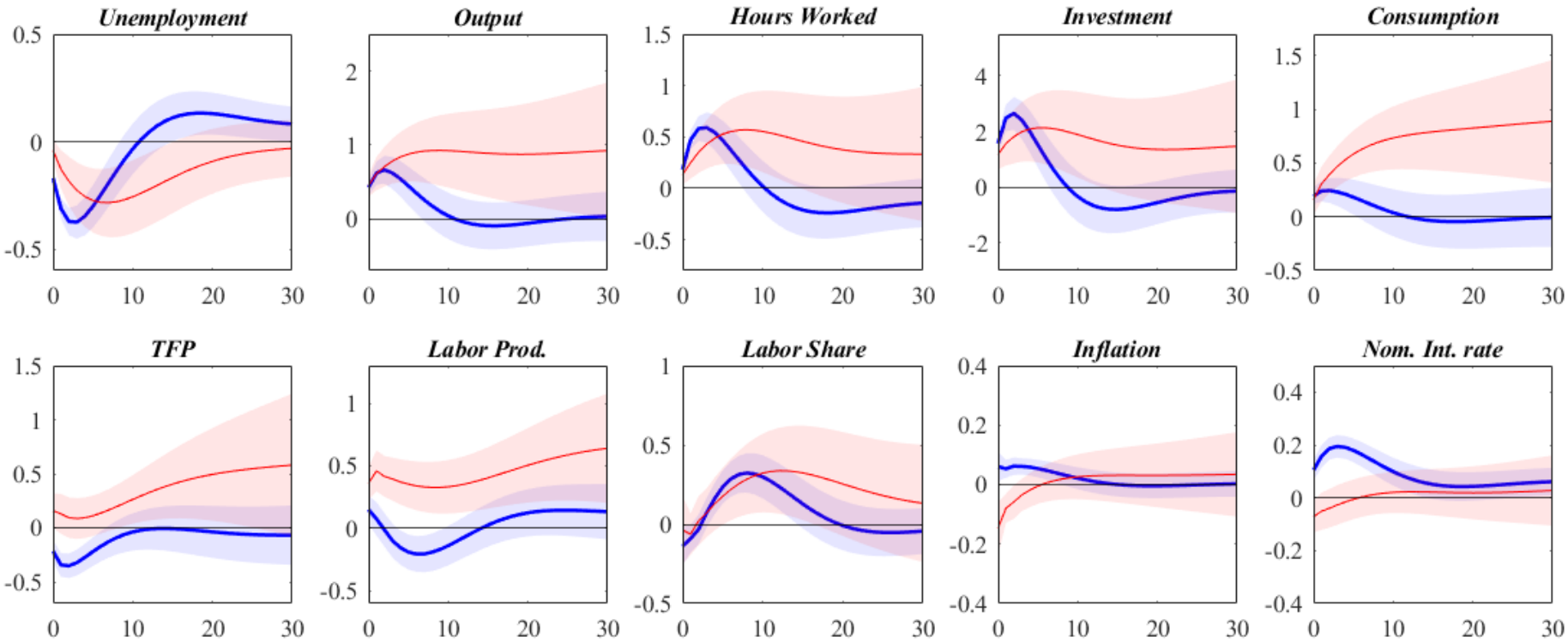
Notes: the table compares the median of the variance contribution of the MBC and the MCBC and, in brackets, their $16th$ and $84th$ percentiles. The results are obtained with bootstrap.

Continued...

- Overall, the responses of the MBC and the MCBC are almost indistinguishable, suggesting that the uncommon component is basically noise.
- We observe a pro-cyclical response to the M(C)BC of investment, output, consumption, hours worked, and unemployment. However, the response of inflation [labor productivity] is positive at all [most] horizons, thus apparently precluding an interpretation as either a traditional demand shock or a supply shock.
- A possible reason why the MCBC has not a straightforward economic interpretation may be that the atheoretical identification scheme picks up a linear combination of two structural shocks, e.g. a demand shock and productivity shock, both of them having a non negligible role in generating the business cycle. Hence, we identify the *Main Common Transitory Business Cycle shock* (MCTBC) as the shock that maximizes the cyclical variability of the common-transitory component of unemployment.

Continued...

Figure: Comparison between CP and MCTBC impulse responses



Notes: the red solid line is the point estimation of the responses to the *common permanent shock*, the blue solid line is the point estimation of the responses to the *main common transitory business cycle shock* whereas the shaded areas are the 1stde confidence bands obtained with bootstrap

Continued...

Table: Measures of contribution of CP and MCTBC

	Short-run		Long-run	
Variable	CP	MTBC	CP	MTBC
Unemployment	17.2[3.3, 47.5]	70.9[44.1, 86.3]	63.8[7.4, 90.7]	14.4[3.0, 58.0]
Output	22.6[4.6, 55.9]	60.7[31.4, 81.2]	85.3[23.9, 97.0]	1.8[0.3, 26.1]
HoursWorked	18.9[4.2, 48.7]	66.7[39.9, 83.6]	73.8[14.6, 95.3]	7.4[1.3, 34.5]
Investment	18.0[3.6, 46.7]	68.1[39.6, 85.0]	82.9[19.2, 96.1]	4.0[0.7, 27.3]
Consumption	21.1[5.8, 53.7]	38.7[13.9, 58.9]	79.8[24.6, 95.4]	2.1[0.2, 24.8]
TFP	12.9[3.3, 40.7]	46.4[17.3, 71.7]	79.5[14.2, 97.2]	2.6[0.2, 29.2]
LaborProd.	20.4[4.4, 46.7]	32.5[16.9, 49.0]	74.1[16.1, 94.5]	4.9[0.5, 37.0]
LaborShare	11.3[3.4, 29.1]	39.2[21.5, 55.1]	72.6[13.0, 92.5]	5.9[1.1, 35.5]
Inflation	23.5[6.4, 57.4]	18.3[5.6, 35.9]	73.9[13.9, 95.7]	3.0[0.3, 24.1]
Nom.Int.rate	12.1[2.6, 34.7]	58.2[32.2, 75.4]	63.9[8.0, 92.4]	8.7[1.3, 51.3]

Notes: the table compares the median of the variance contribution of the CP and MCTBC and, in square brackets, their $16th$ and $84th$ percentiles. The results are obtained with bootstrap techniques.

Conclusions

- In order to achieve dimension reduction for cointegrated medium-VARs, we consider a particular reduced-rank structure on the VECM. The resulting model, i.e. the VECIM, explicitly shows the relation between the variables and a smaller number of observable factors and is a generalization of the MAI Reinsel (1983) to cointegrated time series.
- We provide a switching algorithm for ML estimation of the VECIM parameters, and we show by a MC study that the new method outperforms the VECM when an index structure is present in the data.
- Finally, we apply the VECIM to a set of key US aggregate variables to identify the shock that drives the business cycle.

References I

Angeletos, G. M., Collard, F., and Dellas, H. (2020). Business-cycle anatomy. *American Economic Review*, 110(10):3030–70.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly journal of economics*, 120(1):387–422.

Boswijk, H. (1995). Identifiability of cointegrated systems. *Tinbergen Institute discussion paper*, (TI 95-78).

Bykhovskaya, A. and Gorin, V. (2022). Cointegration in large VARs. *The Annals of Statistics*, 50(3):1593–1617.

Carriero, A., Corsello, F., and Marcellino, M. G. (2022). The global component of inflation volatility. *Journal of Applied Econometrics*, forthcoming.

References II

Carriero, A., Kapetanios, G., and Marcellino, M. (2016). Structural analysis with multivariate autoregressive index models. *Journal of Econometrics*, 192(2):332–348.

Cavaliere, G., De Angelis, L., Rahbek, A., and Taylor, A. (2015). A comparison of sequential and information-based methods for determining the co-integration rank in heteroskedastic VAR models. *Oxford Bulletin of Economics and Statistics*, 77(1):106–128.

Cavaliere, G., De Angelis, L., Rahbek, A., and Taylor, A. (2018). Determining the cointegration rank in heteroskedastic VAR models of unknown order. *Econometric Theory*, 34(2):349–382.

Centoni, M. and Cubadda, G. (2003). Measuring the business cycle effects of permanent and transitory shocks in cointegrated time series. *Economics Letters*, 80(1):45–51.

References III

Cubadda, G. and Guardabascio, B. (2019). Representation, estimation and forecasting of the multivariate index-augmented autoregressive model. *International Journal of Forecasting*, 35(1):67–79.

Cubadda, G., Guardabascio, B., and Hecq, A. (2017). A vector heterogeneous autoregressive index model for realized volatility measures. *International Journal of Forecasting*, 33(2):337–344.

Cubadda, G. and Hecq, A. (2022). Dimension reduction for high dimensional vector autoregressive models. *Oxford Bulletin of Economics and Statistics*, forthcoming.

Engle, R. and Granger, C. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, pages 251–276.

Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco.

Forni, M., Gambetti, L., Marco, M., and Sala, L. (2020). Common component structural VARs. *Available at SSRN 3753929*.

References IV

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2009). Opening the black box: structural factor models with large cross sections. *Econometric Theory*, 25(5):1319–1347.

Giannone, D., Reichlin, L., and M., L. (2008). Explaining the great moderation: It is not the shocks. *Journal of the European Economic Association*, 6(2/3):621–633.

Gonzalo, J. and Pitarakis, J. (1999). Dimensionality effect in cointegration analysis. *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive WJ Granger*, pages 212–229.

Ho, M. and Sorensen, B. (1996). Finding cointegration rank in high dimensional systems using the johansen test: an illustration using data based monte carlo simulations. *Review of Economics and Statistics,*, 78(4):726–732.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press.

References V

Koop, G. M. (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, 28(2):177–203.

Lippi, M. (2018). Frequency-domain approach in high-dimensional dynamic factor models. *Oxford Research Encyclopedia of Economics and Finance*.

Lippi, M. (2019). Time-domain approach in high-dimensional dynamic factor models. *Oxford Research Encyclopedia of Economics and Finance*.

Lippi, M. (2021). Validating DSGE models with SVARs and high-dimensional dynamic factor models. *Econometric Theory*, pages 1–19.

Nielsen, B. (2006). Order determination in general vector autoregressions. *IMS Lecture Notes - Monograph Series*, 52:93–112.

Onatski, A. and Wang, C. (2018). Alternative asymptotics for cointegration tests in large VARs. *Econometrica*, 86(4):1465–1478.

References VI

Reinsel, G. (1983). Some results on multivariate autoregressive index models. *Biometrika*, 70(1):145–156.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.

Stock, J. and Watson, M. (2016). Dynamic Factor Models, Factor-Augmented Vector Autoregressions and Structural Vector Autoregressions in Macroeconomics. In Taylor, J. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2A, pages 415–525. North Holland.