CEIS Tor Vergata Research Papers

Università di Roma Tor Vergata

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 21, Issue 2, No. 556 – April 2023

The Vector Error Correction Index Model: Representation, Estimation and Identification

Gianluca Cubadda and Marco Mazzali

The Vector Error Correction Index Model: Representation, Estimation and Identification*

Gianluca Cubadda†
Università di Roma "Tor Vergata"

Marco Mazzali‡
Università di Roma "Tor Vergata"

March 24, 2023

Abstract

This paper extends the multivariate index autoregressive model by Reinsel (1983) to the case of cointegrated time series of order $(1,1)$. In this new modelling, namely the Vector Error-Correction Index Model (VECIM), the first differences of series are driven by some linear combinations of the variables, namely the indexes. When the indexes are significantly fewer than the variables, the VECIM achieves a substantial dimension reduction w.r.t. the Vector Error Correction Model. We show that the VECIM allows one to decompose the reduced form errors into sets of common and uncommon shocks, and that the former can be further decomposed into permanent and transitory shocks. Moreover, we offer a switching algorithm for optimal estimation of the VECIM. Finally, we document the practical value of the proposed approach by both simulations and an empirical application, where we search for the shocks that drive the aggregate fluctuations at different frequency bands in the US.

Keywords: Vector autoregressive models, multivariate autoregressive index model, cointegration, reduced-rank regression, dimension reduction, main business cycle shock.

*Gianluca Cubadda dedicates this work to the memory of Professor Stefano Fachin, esteemed colleague and dear friend who sadly passed away in February 2023. Previous versions of this paper were presented at the workshop on Dimensionality Reduction and Inference in High-Dimensional Time Series at Maastricht University, the 8th International Conference on Time Series and Forecasting in Gran Canaria, the 2nd Workshop on Time Series Methods for Official Statistics in Paris, the 16th International Conference on Computational and Financial Econometrics in London and an online seminar at the Center for Econometrics and Business Analytics in St. Petersburg. We thank the participants, as well as Giuseppe Cavaliere, Tomás del Barrio Castro, Alain Hecq, Martin Wagner, and two anonymous referees for helpful comments and suggestions. Gianluca Cubadda gratefully acknowledges support by MIUR (PRIN 2020) under the 2020WX9AC7 grant. The usual disclaimers apply.

†Universita' di Roma "Tor Vergata", Centre for Economic and International Studies, Via Columbia 2, 00133 Roma, Italy. Email: gianluca.cubadda@uniroma2.it.

‡Universita' di Roma "Tor Vergata", Dipartimento di Economia e Finanza, Via Columbia 2, 00133 Roma, Italy. Email: marco.mazzali@uniroma2.it.

1 Introduction

Since the seminal paper of Sims (1980), the Vector Autoregressive Model (VAR) has been used largely for forecasting, conducting structural analyses, as well as for investigating the existence of long-term equilibria among economic variables (Granger, 1981; Engle and Granger, 1987). However, the number of parameters to be estimated dramatically increases with the square of the number of variables that the econometrician wishes to include in the analysis. The literature generally addresses the issue as *curse of dimensionality*. Hence, unless we impose some restrictions in order to reduce the number of free parameters, classical VARs are not a viable option to handle medium to large information sets.

The most successful alternatives to the classical VAR in empirical applications are the Dynamic Factor Model (DFM), the Bayesian VAR (BVAR) and, lately, sparse VARs. The attractive feature of DFMs is the assumption that few latent factors drive the entire economy; see *e.g.* Stock and Watson (2016), Lippi et al. (2022), and the references therein. BVARs (see *i.a.* Bańbura et al. (2010), Giannone et al. (2015), Koop (2013)) and sparse VARs (see *i.a.* Hsu et al. (2008), Kock and Callot (2015), Hecq et al. (2021)) rely on different shrinkage parameter approaches to handle medium-large VARs and are commonly used in empirical applications.

Although most macroeconomic indicators are non-stationary, the literature on DFMs mainly focused on stationary data, typically obtained by differencing variables before estimation. However, this practice implicitly assumes that all shocks have permanent effects on the levels of variables, whereas the macroeconomic literature agrees on the fact that some shocks are permanent (*e.g.* technology shocks) and are the source of common trends, while some others are transitory (*e.g.* demand shocks) and generate fluctuations around the trends. However, only a relatively small part of the literature on large systems focused on cointegration. In the DFM framework, notable exceptions are *i.a.* Bai and Ng (2004), Bai (2004), Zhang et al. (2019), Barigozzi et al. (2021), Barigozzi and Trapani (2022), and Casoli and Lucchetti (2022). Even in the fields of regularization methods, contributions on cointegration are relatively scarce; see *e.g.* Diniz et al. (2020), Smeekes and Wijler (2020) and the references therein.

In the classical VAR framework, it is well known that the usual Maximum Likelihood (ML) procedure for cointegration analysis (Johansen, 1995) does not work properly when the number of variables n is relatively large w.r.t. the sample size T. In particular, Monte Carlo studies show that the Likelihood Ratio Test (LRT) tends to overreject the null of fewer cointegration relations in favor of the alternative of more of them; see *i.a.* Ho and Sorensen (1996) and Gonzalo and Pitarakis (1999).[1] While small sample corrections or bootstrap procedures (see *i.a.* Reinsel and Ahn (1992) and Cavaliere et al. (1996)) are effective in mitigating the size distortion of the LRT when the sample size is not large (i.e., $T = 50$) and the dimension is small (i.e., $n \leq 5$), Onatski and Wang (2018) proved that the usual LRT statistic does not converge under the null to the usual asymptotic distribution when both n and T diverge. Recently, Bykhovskaya and Gorin (2022) proposed a variant of the LRT for a VAR(1) model and found its limit distribution in a double asymptotics framework.[2]

In this paper, we follow a different route from previous contributions, as we focus on a medium-dimensional framework given the diffuse evidence that no substantial gains in macroeconomic applications are obtained by further increasing the dimension of the VAR. For instance, when the task is to forecast key aggregate indicators, Bańbura et al. (2010) and Koop (2013) show that the gains in forecast accuracy are

[1] For instance, Gonzalo and Pitarakis (1999) report that in a 9−dimensional VAR(1) with one cointegration vector, the correct decision frequency of the 5% level LR test is 27.24 [87.64] with a sample size equal to 150 [400].

[2] Remarkably, both Onatski and Wang (2018) and Bykhovskaya and Gorin (2022) assume that the number of cointegrating relations is fixed as n increases. This is equivalent to assume that the number of common trends grows with n.

small or even negative by increasing the dimension of the BVARs beyond 20, Cubadda and Guardabascio (2019) document that properly restricted VAR models with 10 or 20 variables outperform larger models, while Forni et al. (2019) argue that medium VARs may contain enough information to identify structural shocks. On the technical side, this implies that we do not need to assume that the number of series diverges for inferential purposes.

Recently, there has been renewed interest in the work of Reinsel (1983), namely the Multivariate Autoregressive Index Model (MAI). The MAI is obtained by imposing a particular reduced rank structure on the coefficient matrices of a VAR such that each variable is driven by the lags of a limited number of linear combinations of the variables, which are called the *indexes* and can be considered as "observable" common factors, in the sense that they are identified even for a finite number of series. Hence, the MAI represents a bridge between DFMs and reduced-rank VARs; see *i.a.* Carriero et al. (2011), Cubadda and Hecq (2011), and Bernardini and Cubadda (2015). The MAI has recently been employed and extended in several directions; see *i.a.* Carriero et al. (2016, 2022), Cubadda et al. (2017), Cubadda and Guardabascio (2019), and Cubadda and Hecq (2022a). The present paper is strongly related to this branch of literature and aims to expand it allowing for cointegration among variables. To the best of our knowledge, previous literature related to the MAI has focused so far on VARs with stationary roots only.

In particular, we aim to achieve dimension reduction for medium-dimensional cointegrated VARs by means of 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), which can be seen as a generalization of the MAI to cointegrated time series. In fact, when there is no cointegration the VECIM reduces to a MAI for the first differences of variables.

An interesting property of the VECIM is that the indexes themselves follow a standard Vector Error-Correction Model (VECM) with a dimension smaller than n. Moreover, the VECIM allows one to decompose the reduced form shocks into sets of common and uncommon shocks, and the former can be further decomposed into permanent and transitory shocks. This opens the possibility of identifying structural shocks from specific directions of the reduced-form errors (e.g., those that are common among variables and have permanent effects).

We offer a switching algorithm for ML estimation, and, in light of previous research indicating that information criteria are particularly useful in selection of the VECM specification (see e.g. Gonzalo and Pitarakis (1999), Cavaliere et al. (2015), and Cavaliere et al. (2018)), we rely on informational methods rather than on testing procedures for the specification of the VECIM.

This paper is organized as follows. Section 2 presents the theoretical aspects, in terms of both model representation and statistical inference. In Section 3 a Monte Carlo study evaluates the finite sample properties of the proposed methods and shows that the inference based on the VECIM systematically outperforms the classical ML analysis of the VECM when an index structure exists. Section 4 provides an empirical application to a medium set of time series to assess the practical usefulness of our approach in studying the propagation mechanism of macroeconomic shocks. The exercise is divided into two parts. First, we decompose the variables into two components, one of which is common while the other is uncommon, and the former is further decomposed into a permanent and a transitory component. Based on Centoni and Cubadda (2003), we provide a measure of the contribution of each component at the business cycle frequency band. Second, we follow Angeletos et al. (2020) and identify the shocks that maximizes the cyclical variability of unemployment and its common component. Moreover, we decompose the main common business cycle driver into permanent and transitory shocks. Finally, Section 5 concludes.

2 Theory

In this Section we first present the derivation and the properties of the proposed modelling, then we discuss statistical inference for all possible specifications.

2.1 Representation theory

Let us assume that the n-vector time series Y_t is generated by the following VAR(p) model

$$Y_t = \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t, \tag{1}$$

where $t = 1, \ldots, T$, Φ_j is an $n \times n$ matrix for $j = 1, \ldots, p$ with $\Phi_p \neq 0$ such that the roots of det $\left(I_n - \sum_{j=1}^{p} \Phi_j z^j\right)$ are equal to 1 or larger than 1 in modulus, ε_t is an n-vector of errors with $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{E}(\varepsilon_t \varepsilon_t') = \Omega$, a positive definite matrix, finite fourth moments, $\mathrm{E}(\varepsilon_t | \mathcal{F}_{t-1}) = 0$ and $\mathcal{F}_t$ is the natural filtration of the process Y_t. For simplicity, we assume that deterministic elements are absent.

We start by rewriting model (1) in the following multivariate augmented Dickey-Fuller representation

$$\Delta Y_t = \Pi_0 Y_{t-1} + \sum_{j=1}^{p-1} \Pi_j \Delta Y_{t-j} + \varepsilon_t, \tag{2}$$

where $\Delta = (1 - L)$, L is the lag operator, $\Pi_0 = \sum_{j=1}^{p} \Phi_j - I_n$, and $\Pi_j = -\sum_{i>j} \Phi_i$ for $j = 1, \ldots, p-1$.

Moreover, we assume that series Y_t follow the Vector Error-Correction Model (VECM)

$$\Delta Y_t = \alpha_0 \beta' Y_{t-1} + \sum_{j=1}^{p-1} \Pi_j \Delta Y_{t-j} + \varepsilon_t, \tag{3}$$

where α_0 and β are full-rank $n \times r$ ($r < n$) matrices such that $\alpha_0 \beta' = \Pi_0$, $\Pi_j = -\sum_{i>j} \Phi_i$ for $j = 1, \ldots, p-1$, ${\alpha_0}'_{\perp} \bar{\Pi} \beta_{\perp}$ is non-singular, and $\bar{\Pi} = I_n - \sum_{j=1}^{p-1} \Pi_j$. Under such assumptions, it is well known that elements of Y_t are individually, at most, $I(1)$ and that they are jointly cointegrated or order 1, in the sense that $\beta' Y_{t-1}$ is $I(0)$, see *i.a.* Johansen (1995) and the references therein.

To possibly reduce the number of parameters in the VECM, we follow Reinsel (1983) and take the following assumptions:

Assumption 1 *For* $\Pi = [\Pi_1', \ldots, \Pi_{p-1}']'$ *it holds that* $\Pi = A\omega'$, *where* ω *is a full-rank* $n \times q$ *matrix with* $q < n$ *and* A *is a full-rank* $n(p-1) \times q$ *matrix.*

Assumption 2 *It holds that* $\beta = \omega\gamma$, *where* γ *is a full-rank* $q \times r$ *matrix with* $q \geq r$.

Under Assumptions 2.1-2.2, the dynamics of the system are endowed with the following index structure:

$$\Delta Y_t = \alpha_0 \gamma' \omega' Y_{t-1} + \sum_{j=1}^{p-1} \alpha_j \omega' \Delta Y_{t-j} + \varepsilon_t, \tag{4}$$

where $A = [\alpha_1', ..., \alpha_{p-1}']'$ and α_j is an $n \times q$ matrix for $j = 1, ..., p-1$.

We call (4) the Vector Error Correction Index Model (VECIM) and label the variables $f_t = \omega' Y_t$ as the indexes. The interpretation of VECIM is that there is a limited number of channels (q over n) through which the first differences ΔY_t are influenced by information coming from the past. This is in line with the traditional view that few shocks are responsible for aggregate fluctuations in the economy. Note that 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.[3]

It is easy to see that VECIM has the following implication for the VAR representation:

$$\Delta Y_t = \sum_{j=1}^{p} \theta_j \omega' Y_{t-j} + \varepsilon_t, \tag{5}$$

where $\theta_{j+1} = \alpha_{j+1} - \alpha_j$ for $j = 1, \ldots, p-1$ with $\alpha_p = 0$, and $\theta_1 = \alpha_0\gamma' - \sum_{j=2}^{p} \theta_j$. In view of Equation (5), we observe that the index structure in (4) prevents a MAI for levels Y_t.

A pleasant property of VAR models with an index structure is that variables f_t have the same dynamic structure as series Y_t (see Carriero et al. (2016) and Cubadda et al. (2017)) and VECIM is no exception. Indeed, premultiplying both sides of Equation (4) by ω' one obtains

$$\Delta f_t = \underline{\alpha}_0 \gamma' f_{t-1} + \sum_{j=1}^{p-1} \underline{\alpha}_j \Delta f_{t-j} + \epsilon_t, \tag{6}$$

where $\underline{\alpha}_0 = \omega'\alpha_0$, $\underline{\alpha}_j = \omega'\alpha_j$ for $j = 1, \ldots, p-1$, and $\epsilon_t = \omega'\varepsilon_t$. Equation (6) tells us that the indexes f_t follow a VECM model with a cointegration matrix equal to γ. Remarkably, when $r = 0$ series ΔY_t are generated by a stationary MAI, whereas in the opposite case $r = q$ series Y_t follow a VECIM where the same linear combinations that stationarize variables Y_t are those that convey information from the past in the first differences ΔY_t.

As noted by Carriero et al. (2016) and Cubadda and Guardabascio (2019), the index structure in VAR models has interesting implications for the shock propagation mechanism. In particular, it is well known that the first differences of cointegrated time series admit the following Wold representation:

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

where $\Psi(L)\left[(1-L)(I_n - \sum_{j=1}^{p-1} \alpha_j \omega' L^j) - \alpha_0 \gamma' \omega' L\right] = \Delta$, $\Psi(1) = \beta_\perp \left(\alpha_{0\perp}' \bar{\Pi} \beta_\perp\right)^{-1} \alpha_{0\perp}'$, see *e.g.* Johansen (1995). Hence, 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

$$Y_t = \chi_t + \iota_t, \tag{7}$$

where

$$\begin{aligned} \chi_t &= C^*(L)\epsilon_t, \\ \iota_t &= \Psi^*(L)\omega_\perp(\omega_\perp'\Omega^{-1}\omega_\perp)^{-1}\xi_t, \end{aligned} \tag{8}$$

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

Since $\mathrm{E}(\epsilon_t'\xi_t) = 0$, the components χ_t and ι_t are not correlated at all the lags and leads. Given that χ_t has the same innovations as the indexes f_t, it is legitimate to interpret the former as the common component

[3]Since matrices ω and γ, once identified through normalizing restrictions, respectively have $q(n-q)$ and $r(q-r)$ unknown elements, the VECIM has $r(n+q-r) + q(np-q)$ free parameters.

of the series Y_t. Hence, one can use the methods adopted in the structural DFM literature (see, *i.a.*, Forni et al. (2009, 2020)) to recover structural shocks from the reduced-form common shocks ϵ_t.

Moreover, when $0 < r < q$ it is possible to further decompose the reduced-form common shocks ϵ_t into a component having permanent effects on Y_t and another one having transitory effects only.[4] Indeed, by putting between $C(L)$ and ϵ_t in Equation (8) 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

$$\chi_t = \pi_t + \tau_t, \tag{9}$$

where

$$\pi_t = \underbrace{C^*(L)\Sigma\underline{\alpha}_{0\perp}(\underline{\alpha}'_{0\perp}\Sigma\underline{\alpha}_{0\perp})^{-1}}_{P^*(L)}\underbrace{\underline{\alpha}'_{0\perp}\epsilon_t}_{u_t},$$
$$\tau_t = \underbrace{C^*(L)\underline{\alpha}_0(\underline{\alpha}'_0\Sigma^{-1}\underline{\alpha}_0)^{-1}}_{T^*(L)}\underbrace{\underline{\alpha}'_0\Sigma^{-1}\epsilon_t}_{\eta_t}$$

Since u_t are the innovations of the common trends in the multivariate Beveridge-Nelson decomposition of the indexes f_t (see, e.g., Johansen (1995)) and given that $\mathrm{E}(u'_t\eta_t) = 0$, 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, which are not correlated with each other in any lag and lead.[5]

Remarkably, the VECIM allows us to conduct structural analysis by taking advantage of the features of both the DFM, i.e. isolating shocks that are common among variables, and of the VECM, i.e. disentangling shocks having either transitory or permanent effects. For instance, one may obtain the structural common permanents shocks as $\upsilon_t = C^{-1}Du_t$ and the impulse response functions from $\tilde{P}^*(L) = P^*(L)D^{-1}C$, where D is the matrix formed by the first $s = q - r$ rows of $P^*(0)$ and C is a lower triangular matrix such that $CC' = D\Sigma D'$. Since the first s rows of $\tilde{P}^*(0)$ form a lower triangular matrix, the usual interpretation of structural shocks υ_t applies as long as the m ($m \leq s$) variables of interest are placed and properly ordered in the first m elements of Y_t.

Furthermore, it is possible to measure the variability that is explained by some specific direction, or all directions, of each component of the shocks at a given frequency band. Indeed, being the components in (7) and (9) not cross-correlated at any lag and lead, we can decompose the spectral density of the series Y_t as follows:

$$F(\lambda) = F_\pi(\lambda) + F_\tau(\lambda) + F_\iota(\lambda) \tag{10}$$

where

$$F(\lambda) = (1/2\pi)^{-1}\Psi^*(z)\Omega\Psi^*(z^{-1})',$$
$$F_\pi(\lambda) = (1/2\pi)^{-1}P^*(z)\underline{\alpha}'_{0\perp}\Sigma\underline{\alpha}_{0\perp}P^*(z^{-1})',$$
$$F_\tau(\lambda) = (1/2\pi)^{-1}T^*(z)\underline{\alpha}'_0\Sigma^{-1}\underline{\alpha}_0T^*(z^{-1})',$$

$z = \exp(-i\lambda)$, and $\lambda \in (0, \pi]$.

[4]Notice that when $r = q$ the common component χ_t is $I(0)$, whereas in the opposite case $r = 0$ we have that $\chi_t \sim I(1)$. In both cases, no further decompositions of χ_t in permanent-transitory components are possible.

[5]This distinctive feature of the decomposition in (9) is not shared by other popular trend-cycle decompositions such as Kasa (1992), Gonzalo and Granger (1995) and Zhang et al. (2019) .

Since the process Y_t has some roots equal to one, its spectral density matrix is unbounded at the zero frequency. Hence, $F(\lambda)$ is strictly speaking a pseudospectral density matrix, but, for the sake of simplicity, we will omit the adjective pseudo in the following. If it is of interest to analyze the spectral density matrix of the first differences ΔY_t rather than $F(\lambda)$, it suffices to substitute $\Psi^*(L)$ with $\Psi(L)$ in the formulae underlying the decompositions (7), (9), and (10).

The decomposition (10) enables us not only to measure the effects of each component shock in a given frequency band as in Centoni and Cubadda (2003) but also to do the same for a specific direction of such shocks. For instance, the contribution of the first of the structural common-permanent shocks υ_t to the spectral density of the $k-$th element of the series Y_t in the frequency band $[\lambda_a, \lambda_b]$, with $0 < \lambda_a \leq \lambda_b \leq \pi$, is given by

$$\int_{\lambda_a}^{\lambda_b} e_k' \tilde{p}_1^*(z) \tilde{p}_1^*(z^{-1})' e_k,$$

where $\tilde{p}_1^*(z)$ is the first column of the matrix $\tilde{P}^*(L)$, and e_k is an $n-$vector with unity as its $k-$th element and zeroes elsewhere.

Remark 1 *An alternative modelling to the VECIM consists in taking Assumption 2.1 only without introducing any constraint on the cointegration matrix β. It is easy to see that this is equivalent to assume that all the VAR coefficient matrices but Φ_1 have an index structure. The resulting formulation reads*

$$\Delta Y_t = \alpha_0 \beta' Y_{t-1} + \sum_{j=1}^{p-1} \alpha_j \omega' \Delta Y_{t-j} + \varepsilon_t, \tag{11}$$

Although the above model has the possible advantage of not requiring $q \geq r$, we favor the VECIM for a twofold reason. First, it is easy that the indexes $\omega' Y_t$ in Equation (11) do not follow a VECM model as in the case of Equation (4). Second, and related to the first point, if the data are generated by the model in (11), it is generally not possible to separate the common component χ_t into permanent-transitory components, thus compromising a key feature of the proposed methodology.

Remark 2 *Another alternative to the VECIM consists in assuming a MAI for the levels of the series, i.e $\Phi = \bar{A}\varpi'$, where $\Phi = [\Phi_1', \ldots, \Phi_p']'$ and ϖ is a full-rank $n \times q$ matrix with $q < n$ and $\bar{A}$ is a full-rank $np \times q$ matrix. It is easy to see that under such assumption we have $\Pi \varpi_\perp = 0$ and $\alpha_0 \beta' \varpi_\perp = -\varpi_\perp$, which is a special case of model (11). Hence, the same considerations as in the previous remark apply.*

Remark 3 *A closely related modelling to VECIM is the Reduced-Rank VECM (Vahid and Engle, 1993) that reads*

$$\Delta Y_t = \varphi \vartheta_0' \beta' Y_{t-1} + \varphi \sum_{j=1}^{p-1} \vartheta_j \Delta Y_{t-j} + \varepsilon_t, \tag{12}$$

where φ is a full-rank $n \times q$ matrix with $r \leq q < n$ and $[\vartheta_0', \vartheta_1', \ldots, \vartheta_{p-1}']'$ is a full-rank $[n(p-1)+r] \times q$ matrix. Due to its interpretation in terms of common trends and common cycles in the multivariate Beveridge-Nelson decomposition, model (12) has extensively been applied and generalized in various directions, see e.g. Cubadda and Hecq (2022b) and the references therein. Moreover, it implies that the marginal processes of series Y_t follow parsimonious univariate models, thus solving the so-called autoregressivity paradox (Cubadda et al., 2009). Although models (4) and (12) have a similar structure and the same number of parameters, we stress that the latter does not allow to disentangle shocks that are common among the variables as the former does.

Remark 4 *A contribution that has close analogies to the VECIM in terms of representation theory is Barigozzi et al. (2021), who propose a large DFM with I(1) factors that can be cointegrated and are driven by a number of shocks that is smaller than their dimension. These features are shared by the common component* χ_t *in* (7). *However, unlike Barigozzi et al. (2021), it is not necessary to let* n *diverge to infinity to identify and estimate the common component* χ_t *in our framework.*

2.2 Estimation and model identification

Based on Boswijk (1995) and Cubadda et al. (2017), we offer some switching algorithms that have the property of increasing the Gaussian likelihood of model (4) in each step. In detail, when $0 < r < q$, the procedure goes as follows:

1. Given (initial) estimates of both γ and ω, maximize the conditional Gaussian likelihood $\mathcal{L}(A^{\dagger}, \Omega|\gamma, \omega)$ by estimating $A^{\dagger} = [\alpha_0', A']'$ and Ω with OLS on Equation (4).

2. Premultiply by $\Omega^{-1/2}$ and apply the Vec operator to both the sides of Equation (4), then use the property $\text{Vec}(ABC) = (C' \otimes A)\text{Vec}(B)$ to get

$$\Omega^{-1/2}\Delta Y_t = \left(Y_{t-1}' \otimes \Omega^{-1/2}\alpha_0\gamma' + \sum_{j=1}^{p-1} \Delta Y_{t-j}' \otimes \Omega^{-1/2}\alpha_j \right) \text{Vec}(\omega') + \Omega^{-1/2}\varepsilon_t \tag{13}$$

   Given the previously obtained estimates of $A^{\dagger}$, γ, and Ω, maximize $\mathcal{L}(\omega|A^{\dagger}, \gamma, \Omega)$ by estimating $\text{Vec}(\omega')$ with OLS on Equation (13).

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

$$S_{11}^{-1} S_{10} S_{00}^{-1} S_{01} \tag{14}$$

   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 $[\Delta Y_{t-1}'\omega, \ldots, \Delta Y_{t-p+1}'\omega']'$.

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

When $r = 0$, step 3 is clearly not needed and step 1 and 2 must be modified as follows:

1.1 Given (initial) estimates of ω, maximize $\mathcal{L}(A, \Omega|\omega)$ by estimating A and Ω with OLS on the following model

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

2.1 Given the previously obtained estimates of A and Ω, maximize $\mathcal{L}(\omega|A, \Omega)$ by estimating $\text{Vec}(\omega')$ with OLS on the following model

$$\Omega^{-1/2}\Delta Y_t = \left(\sum_{j=1}^{p-1} \Delta Y_{t-j}' \otimes \Omega^{-1/2}\alpha_j \right) \text{Vec}(\omega') + \Omega^{-1/2}\varepsilon_t \tag{15}$$

Finally, when $r = q$, we can assume without loss of generality that $\gamma = I_q$. Then step 3 is again not needed, whereas steps 1 and 2 must be modified as follows

1.3 Given (initial) estimates of ω, maximize $\mathcal{L}(A^\dagger, \Omega|\omega)$ by estimating $A^\dagger$ and Ω with OLS on the following model

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

2.3 Given the previously obtained estimates of $A^\dagger$ and Ω, maximize $\mathcal{L}(\omega|A^\dagger, \Omega)$ by estimating $\mathrm{Vec}(\omega')$ with OLS on the following model

$$\Omega^{-1/2}\Delta Y_t = \left(Y'_{t-1} \otimes \Omega^{-1/2}\alpha_0 + \sum_{j=1}^{p-1} \Delta Y'_{t-j} \otimes \Omega^{-1/2}\alpha_j \right) \mathrm{Vec}(\omega') + \Omega^{-1/2}\varepsilon_t \tag{16}$$

As argued by Cubadda and Guardabascio (2019), switching algorithms have several advantages over Newton-type optimization methods, such as computational simplicity, no need for normalization conditions on ω and γ, explicit optimization at each step, and ease of application of regularization schemes or linear restrictions on parameters. Furthermore, when the switching algorithm is initialized with consistent estimates and is iterated sufficiently often, the resulting estimator is asymptotically equivalent to the ML estimator (MLE) being numerically approximated (Hautsch et al., 2022). Therefore, a proper choice of initial values for ω and γ is crucial to boost numerical and statistical convergence. For the former, we rely on the right-singular vectors that correspond to the q largest singular values of the matrix $\alpha_0\beta' + \sum_{j=1}^{p-1}\Gamma_j$ having estimated the coefficient matrices of the VECM with the usual Johansen procedure. For the latter, we suggest using the relation $\gamma = (\omega'\omega)^{-1}\omega'\beta$ having fixed ω to its initial value.

In order to specify the values for p, r, and q, two alternative approaches are in principle viable. The first one relies on sequential LR testing in order to fix initially p, then r conditionally on p and finally q conditionally on p and r. The second approach consists in simultaneously establishing the triple (p, r, q) by means of informational methods. In particular, model (4) is estimated for all plausible values of (p, r, q) and the triple that minimizes a certain information criterion (IC) is selected. Most common IC may be used, *i.e.* those named after Akaike (AIC), Hannan-Quinn (HQIC) and the Bayesian IC (BIC). In line with previous contributions indicating that consistent IC are particularly effective in specification of cointegrated VAR models (see e.g. Gonzalo and Pitarakis (1999), Cavaliere et al. (2015), and Cavaliere et al. (2018)), we opt for determining the triple (p, r, q) in one single search.[6]

Remark 5 *As correctly pointed out by the referees, economic theory often dictates zero restrictions on the cointegration matrix β. Given the index structure of the VECIM, this requires the existence of r_1 ($r_1 \leq r$) cointegrating vectors with the following form*

$$\beta'_1 = \underset{r_1 \times q_1}{\gamma'_1} [\underset{q_1 \times n_1}{\omega'_1}, \underset{q_1 \times (n-n_1)}{0}]$$

where $r_1 \leq q_1 \leq n_1 < n$. If the zero restrictions on both ω and γ are overidentifying,[7] *constrained estimation of ω simply requires to impose the appropriate zeros in OLS estimation of model (13), whereas for the restrictions on γ it is convenient to modify step 3 of the switching algorithm as follows*

[6]Remarkably, BIC and HQIC remain consistent for r even under forms of heteroskedasticity that invalidate the limit distribution of the LR test, and joint IC-based estimation of p and r outperforms in finite samples sequential procedures based on either IC or tests (Cavaliere et al. (2018)).

[7]Notice that just-identifying restrictions such as $\omega' = [I_q, \nu']$, where ν is a $(n-q) \times q$ matrix, can be easily imposed in estimation, see p. 339 of Cubadda et al. (2017),.

3.1 Proceed as in step 2 but with a different factorization of the matrix $\Omega^{-1/2}\alpha_0\gamma'\omega'$ *when applying* $\mathrm{Vec}(ABC)$ *to get*

$$\Omega^{-1/2}\Delta Y_t - \left(\sum_{j=1}^{p-1}\Delta Y'_{t-j}\otimes\Omega^{-1/2}\alpha_j\right)\mathrm{Vec}(\omega') \tag{17}$$
$$= \left(Y'_{t-1}\omega\otimes\Omega^{-1/2}\alpha_0\right)\mathrm{Vec}(\gamma') + \Omega^{-1/2}\varepsilon_t$$

Given the previously obtained estimates of α, ω, *and* Ω, *maximize* $\mathcal{L}(\gamma|\alpha,\omega,\Omega)$ *by estimating* $\mathrm{Vec}(\gamma')$ *with OLS on Equation (17).*[8]

Then it is straightforward to impose zero restrictions on γ *in the estimation of the model (17). Restricted VECIM specifications can be compared with the unrestricted one using IC.*

Remark 6 *As a referee stressed, if* ω *were known, Assumption 2.2 would correspond to the case of the same restrictions on all cointegration vectors in Johansen (1995). Therefore, the method in step 3 would provide the MLE of* γ *(then the MLE of* $\beta=\omega\gamma$*), which would converge at rate* T *and its limit distribution would be mixed normal. However, when* ω *is unknown, since it is a coefficient matrix of both* Y_{t-1} *and* $[\Delta Y'_{t-1},\ldots,\Delta Y'_{t-p+1}]'$ *in (4), its MLE converges at the slowest rate of the estimated coefficients that are attached to either* $I(1)$ *or* $I(0)$ *regressors only, i.e.* $\sqrt{T}$ *(see, e.g., Sims et al. (1990)). If one wishes to preserve the asymptotic properties of the unrestricted MLE of* β*, a route to go is first to estimate the cointegration vectors by the usual Johansen procedure, then to estimate* ω *and* $A^{\dagger}$ *conditionally on* β *by the switching algorithm but without step 3 and having substituted (13) in step 2 with the following:*

$$\Omega^{-1/2}\left(\Delta Y_t - \alpha_0\beta' Y_{t-1}\right) = \left(\sum_{j=1}^{p-1}\Delta Y'_{t-j}\otimes\Omega^{-1/2}\alpha_j\right)\mathrm{Vec}(\omega') + \Omega^{-1/2}\varepsilon_t$$

Finally, γ *could be estimated through the relation* $\gamma=(\omega'\omega)^{-1}\omega'\beta$ *having fixed both* ω *and* β *to their estimates. Given the super-consistency of the unrestricted MLE of* β*, it is easy to recognize that this procedure leads to an estimator of the parameters of model (11) that is asymptotically equivalent to the MLE, thus providing inefficient inference for the VECIM parameters.*[9]

3 Monte Carlo analysis

In this Section we perform a Monte Carlo study to evaluate the finite sample performances of the proposed approach. We consider the following n-dimensional cointegrated VAR(3) process

$$Y_t = \sum_{j=1}^{3}\Phi_j Y_{t-j} + \varepsilon_t, \tag{18}$$

where $t=1,\ldots,T$, $\Phi_1=\omega^{+}(\mathrm{diag}(\delta_1)-I_q)\omega'+I_n$, $\omega^{+}=\omega(\omega^{'}\omega)^{-1}$, the elements of ω are generated by independent $\mathrm{U}(-1,1)$ distributions, $\delta_1=[\delta'_{1,1},\delta'_{1,2}]'$, $\delta_{1,1}=2m_1\odot\cos(\lambda_1)+\rho_1$, $\delta_{1,2}=2m_2\odot\cos(\lambda_2)+1_{q-r}$, $\odot$ denotes element-wise multiplication, m_1 is an $r-$vector and m_2 is an $(q-r)-$vector such that $m=[m'_1,m'_2]'=$

[8] We do not suggest to resort to (17) in place of (14) if no restrictions on γ must be imposed. The reason is that in the former γ is estimated conditionally on α,ω,Ω whereas in the latter γ is estimated conditionally on ω only, thus speeding up numerical convergence of the switching algorithm.

[9] On the light of the beforementioned contributions documenting the poor small-sample properties of the Johansen procedure in medium dimensional VARs, which will be confirmed by our Monte Carlo study, we do no suggest such approach.

$1_q 0.7$, λ_1 is an $r-$vector and λ_2 is an $(q-r)-$vector such that $\lambda = [\lambda_1', \lambda_2']'$ is drawn from a $\mathrm{U}_q[\pi/16, \pi/3]$, ρ_1 is an $r-$vector and ρ_2 is an $(q-r)-$vectors such that $\rho = [\rho_1', \rho_2']' = 1_n 0.7$, $\Phi_j = \omega^+ \mathrm{diag}(\delta_j)\omega'$ and $\delta_j = [\delta_{j,1}', \delta_{j,2}']'$ for $j = 2, 3$, $\delta_{2,1} = -[2\rho_1 \odot m_1 \odot \cos(\lambda_1) + m_1 \odot m_1]$, $\delta_{2,2} = -[2m_2 \odot \cos(\lambda_2) + m_2 \odot m_2]$, $\delta_{3,1} = \rho_1 \odot m_1 \odot m_1$, $\delta_{3,2} = m_2 \odot m_2$, and ε_t are i.i.d. $\mathrm{N}_n(0, I_n)$. Notice that the cointegration matrix β is composed of the first r columns of ω when $r > 0$.

Some remarks are in order. Premultiplying both sides of (18) by ω' we see that indexes f_t follow a diagonal VAR(3) with $q - r$ roots equal to 1, and r real inverse roots that are equal to 0.7 as in Gonzalo and Pitarakis (1999), and q pairs of complex conjugate inverse roots with a modulus that is equal to 0.7 and angular frequencies that belong to the business cycle band for quarterly data.[10] Premultiplying both sides of (18) by $\omega_\perp'$ we see that the uncommon component $\omega_\perp' Y_t$ follows instead a multivariate random walk. Hence, the DGP is constructed to mimic a system of $I(1)$ quarterly time series with common trends and common business cycles.

From (18) we simulate systems with $q = 2, 4, 6$ and $r = 0, q/2, q$ for $n = 8, 12, 16$ variables. The number of observations is $T = 240, 480, 720$. We generate $T + 50$ observations, and the first 50 points are used as a burn-in period, the remaining ones for estimation.

The proposed approach is evaluated in combination with the various IC using the following statistics. First, the percentage of correct estimation of the number of indexes q. Second, the percentage of correct estimation of the couple (p, r). Third, the Frobenius distance between the estimated VAR coefficients and the true ones, relative to the Frobenius norm of the true coefficients (RFD). Fourth, the average of the mean square 1-step ahead forecast errors over the n series (AMSFE). For comparative purposes, all the statistics but the first one are also computed for the VECM estimated by the usual Johansen procedure. The results reported below are based on 1000 replications for each combination.

From Table (1) we see that the HQIC outperforms competitors in estimating q but the cases with $n = 8$, for which the BIC performs similarly. Interestingly, the HQIC offers a percentage of correct identification that is equal to or close to 100% in all cases, whereas the AIC systematically overestimates the correct number of indexes.

In Table (2) VECIM and VECM are compared in terms of the percentage of correct identification of the couple (p, r). Since the HQIC performs best by a clear margin, we limit our comments to the results obtained with this IC. We observe that the percentage of correct estimation of the couple (p, r) is systematically higher for VECIM in all cases. For VECM, this percentage is low, often equal to zero, when n is greater than 8. This evidence confirms that the VECM clearly suffers from the dimensionality problem. Even the performance of the VECIM deteriorates with n, in particular when $q = r$. It should be noted that in such cases our DGP imposes q stationary real roots that are not far from 1, which is notoriously a circumstance where all cointegration tests suffer from a lack of power.

Table (3) compares VECIM and VECM in terms of accuracy of estimation measured by the RFD. First, we notice that the HQIC tends to perform best and the BIC stays slightly behind. Second, the VECIM systematically provides more efficient estimates than the VECM, even with the smallest system dimension. This evidence shows that significant improvements in inference can be obtained by exploiting the possible existence of common components in the cointegrated VAR model.

Table (4) reports the ARMSFE of both models. Remarkably, HQIC and BIC perform very similarly in terms of forecast accuracy. Again, the VECIM outperforms the VECM in all cases, although the gap between the two methods tends to decrease as the sample size increases.

[10]Notice that the AR polynomial of the generic index $f_{i,t}$ is $(1 - \rho_i L)(1 - 2m_i \cos(\lambda_i)L + m_i^2 L^2)$ for $i = 1, ..., q$.

Table 1: Percentages of correct estimation of q for VECIM

			T =240			T = 480			T = 720		
N	q	r	AIC	BIC	HQIC	AIC	BIC	HQIC	AIC	BIC	HQIC
8	2	0	38.3	99.6	100.0	46.4	100.0	100.0	45.4	100.0	100.0
		1	41.5	99.1	99.9	53.6	100.0	100.0	58.5	100.0	100.0
		2	47.7	98.7	99.0	63.7	100.0	100.0	66.3	100.0	100.0
	4	0	25.0	99.5	99.9	37.8	100.0	100.0	46.5	100.0	100.0
		2	42.8	98.6	99.9	56.8	100.0	100.0	62.3	100.0	100.0
		4	56.4	99.9	98.4	71.4	100.0	100.0	78.6	100.0	100.0
	6	0	32.7	100.0	99.3	48.9	100.0	100.0	53.5	100.0	100.0
		3	54.8	99.5	99.4	68.2	100.0	99.9	78.1	100.0	100.0
		6	80.0	100.0	100.0	91.2	100.0	100.0	91.2	100.0	100.0
12	2	0	16.8	78.8	100.0	35.0	100.0	100.0	39.2	100.0	100.0
		1	23.5	68.5	99.8	40.4	100.0	100.0	47.2	100.0	100.0
		2	23.3	73.1	99.3	52.3	100.0	100.0	56.8	100.0	100.0
	4	0	7.4	70.1	99.9	25.9	100.0	100.0	32.9	100.0	100.0
		2	18.8	64.0	99.6	43.8	100.0	100.0	49.4	100.0	100.0
		4	26.2	69.4	96.5	54.3	100.0	100.0	63.4	100.0	100.0
	6	0	8.5	77.2	99.9	22.3	100.0	100.0	30.1	100.0	100.0
		3	16.8	65.5	98.9	41.7	100.0	100.0	50.9	100.0	100.0
		6	30.0	77.1	94.9	66.1	100.0	99.8	75.4	100.0	100.0
16	2	0	10.0	52.4	98.2	25.0	96.9	100.0	33.4	100.0	100.0
		1	9.7	44.4	96.3	28.7	94.2	100.0	39.2	100.0	100.0
		2	7.2	34.0	98.2	36.8	96.2	100.0	45.5	100.0	100.0
	4	0	2.5	33.5	95.5	16.4	97.5	100.0	24.9	100.0	100.0
		2	3.5	23.6	94.1	25.1	93.8	100.0	35.8	100.0	100.0
		4	5.1	18.7	90.9	38.7	97.5	99.8	50.2	100.0	100.0
	6	0	1.1	31.0	94.3	12.6	98.0	100.0	15.7	100.0	100.0
		3	2.7	26.6	90.6	22.4	94.8	100.0	38.0	100.0	100.0
		6	4.7	18.6	90.0	43.4	98.4	100.0	59.4	100.0	100.0

Notes: Percentages with which each IC correctly estimates the true number of indexes q.

Furthermore, we report in Table (5) the percentages of correct identification of the number of indexes when the data are generated by a VECM without index structure (*i.e.* $q = n$). For the sake of space, we report the results only for $n = 12$.[11] In almost all cases, the estimates of q are equal to n, confirming that the suggested identification approach leads to correctly specify the model even when there is no common component in the data.

4 Empirical application

In this section, we illustrate the practical value of our methodology. We start by comparing the VECIM and the VECM in modelling ten key aggregate US time series, then we use the decomposition by Centoni and Cubadda (2003) to recover the common components of the variables and to further decompose it into

[11] Results for $n = 8, 16$ are available upon request.

Table 2: Percentages of correct identification of the couple (p, r) for the VECIM (VECM)

			T =240			T = 480			T =720		
N	q	r	AIC	BIC	HQIC	AIC	BIC	HQIC	AIC	BIC	HQIC
8	2	0	29.6 (0.4)	99.4 (0.0)	99.7 (13.5)	32.6 (1.8)	100.0 (3.9)	99.6 (94.4)	33.0 (0.7)	100.0 (83.3)	100.0 (97.7)
		1	40.8 (1.6)	11.2 (0.0)	80.2 (1.1)	47.9 (2.9)	74.3 (0.0)	99.2 (56.8)	52.9 (3.5)	96.7 (2.3)	99.8 (95.8)
		2	55.8 (2.1)	0.1 (0.0)	45.7 (0.0)	67.7 (4.7)	57.9 (0.0)	99.6 (0.9)	70.4 (6.0)	97.1 (0.0)	100.0 (33.3)
	4	0	5.6 (0.1)	99.9 (27.9)	94.3 (74.9)	13.1 (1.1)	100.0 (100.0)	98.9 (94.4)	17.0 (2.3)	100.0 (100.0)	99.3 (97.6)
		2	29.1 (3.6)	3.3 (0.0)	66.6 (32.0)	41.8 (6.5)	55.3 (7.1)	96.7 (87.0)	44.7 (6.7)	93.8 (55.2)	99.4 (94.4)
		4	69.0 (13.5)	0.0 (0.0)	57.7 (0.4)	78.0 (10.8)	78.3 (0.0)	99.9 (78.4)	83.8 (10.2)	99.8 (20.5)	100.0 (91.0)
	6	0	0.9 (0.0)	100.0 (99.7)	82.6 (72.5)	4.1 (1.3)	100.0 (100.0)	98.3 (95.4)	5.4 (0.9)	100.0 (100.0)	96.9 (95.2)
		3	19.3 (6.7)	1.3 (0.7)	54.9 (40.7)	23.0 (6.0)	47.9 (16.6)	94.6 (88.6)	24.5 (7.0)	93.2 (74.6)	95.4 (95.5)
		6	87.8 (30.2)	0.0 (0.0)	84.3 (19.1)	94.7 (28.3)	96.2 (24.4)	100.0 (77.2)	93.5 (27.3)	100.0 (98.2)	100.0 (82.1)
12	2	0	14.7 (0.0)	82.5 (0.0)	99.4 (0.0)	28.6 (0.2)	100.0 (0.0)	100.0 (0.1)	34.8 (0.0)	100.0 (0.0)	100.0 (47.5)
		1	25.6 (0.0)	0.1 (0.0)	42.0 (0.0)	40.7 (0.1)	19.2 (0.0)	93.0 (0.0)	48.0 (0.0)	70.1 (0.0)	99.3 (0.5)
		2	30.0 (0.0)	0.0 (0.0)	3.4 (0.0)	59.7 (0.0)	0.4 (0.0)	86.8 (0.0)	61.8 (0.4)	49.8 (0.0)	99.3 (0.0)
	4	0	2.6 (0.0)	92.9 (0.0)	99.4 (1.2)	16.7 (0.0)	100.0 (0.1)	99.8 (94.5)	18.1 (0.0)	100.0 (52.4)	100.0 (98.8)
		2	24.1 (0.0)	0.0 (0.0)	17.0 (0.0)	38.3 (0.8)	4.4 (0.0)	87.6 (19.4)	42.5 (0.7)	49.1 (0.0)	99.3 (86.2)
		4	44.0 (0.6)	0.0 (0.0)	0.6 (0.0)	65.6 (2.1)	0.0 (0.0)	87.7 (0.0)	69.7 (1.6)	44.9 (0.0)	99.9 (3.7)
	6	0	1.0 (0.0)	96.0 (0.2)	95.4 (44.1)	3.7 (0.0)	100.0 (83.4)	98.6 (95.2)	7.7 (0.0)	100.0 (100.0)	99.9 (99.0)
		3	14.2 (0.3)	0.0 (0.0)	8.9 (0.3)	27.8 (0.5)	1.3 (0.0)	83.7 (46.1)	33.4 (0.9)	33.4 (0.3)	98.7 (86.0)
		6	48.5 (7.8)	0.0 (0.0)	0.6 (0.0)	77.2 (6.9)	0.0 (0.0)	96.5 (1.9)	81.4 (4.8)	70.8 (0.0)	100.0 (87.5)
16	2	0	10.3 (0.0)	37.9 (0.0)	96.9 (0.0)	24.0 (0.0)	100.0 (0.0)	100.0 (0.0)	32.9 (0.0)	100.0 (0.0)	100.0 (0.0)
		1	17.0 (0.0)	0.0 (0.0)	13.5 (0.0)	32.9 (0.0)	1.1 (0.0)	73.6 (0.0)	42.2 (0.0)	27.8 (0.0)	97.0 (0.0)
		2	14.3 (0.0)	0.0 (0.0)	0.0 (0.0)	44.5 (0.0)	0.0 (0.0)	39.6 (0.0)	53.1 (0.0)	1.6 (0.0)	93.7 (0.0)
	4	0	1.3 (0.0)	38.9 (0.0)	99.1 (0.0)	11.8 (0.0)	100.0 (0.0)	100.0 (0.9)	18.4 (0.0)	100.0 (0.0)	100.0 (88.3)
		2	12.8 (0.0)	0.0 (0.0)	2.5 (0.0)	26.0 (0.0)	0.0 (0.0)	56.2 (0.0)	34.6 (0.0)	6.5 (0.0)	92.7 (0.2)
		4	19.9 (0.0)	0.0 (0.0)	0.0 (0.0)	52.2 (0.0)	0.0 (0.0)	31.5 (0.0)	58.3 (0.1)	0.0 (0.0)	94.6 (0.0)
	6	0	0.1 (0.0)	40.9 (0.0)	97.8 (0.0)	4.6 (0.0)	100.0 (0.0)	99.9 (91.6)	6.6 (0.0)	100.0 (0.5)	100.0 (98.8)
		3	9.3 (0.0)	0.0 (0.0)	0.9 (0.0)	22.1 (0.0)	0.0 (0.0)	40.4 (0.0)	31.4 (0.0)	0.8 (0.0)	91.3 (38.0)
		6	29.2 (0.1)	0.0 (0.0)	0.0 (0.0)	59.9 (1.5)	0.0 (0.0)	37.2 (0.0)	68.2 (0.8)	0.0 (0.0)	96.6 (0.0)

Notes: Percentages with which each IC correctly estimates the true couple (p, r) for the VECIM and, in parentheses, for the VECM.

permanent and transitory components, finally we show how the VECIM may contribute to better interpret two shocks that are obtained by applying the identification procedure of Angeletos et al. (2020) to the common permanent-transitory components of the considered variables.

4.1 Comparison with the VECM

In this empirical analysis we use quarterly US macroeconomic data from the FRED database. The sample covers the period between $1955Q1$ and $2019Q4$ and variables are transformed such that they are at most $I(1)$. Specifically, we take the following ten variables: real GDP per capita, real consumption per capita, computed as the sum of non-durable consumption and services, real investment per capita, computed as the sum of investment and durable consumption, hours worked per person, inflation, obtained from the difference of the log of the GDP deflator, the unemployment rate, the nominal interest rate, as measured by the FED Funds rate, TFP, which is the cumulated sum of the utilization-adjusted TFP as computed in Fernald (2014), the non-farm business sector labor productivity and labor share. We take the log of real variables and the first difference of the log of nominal variables. The remaining variables are left unmodified.

As a first step, it is of interest to check whether a VECM or a VECIM fits better to the data by means

Table 3: RFD of the VAR coefficients for the VECIM (VECM)

			T =240			T = 480			T =720		
N	q	r	AIC	BIC	HQIC	AIC	BIC	HQIC	AIC	BIC	HQIC
8	2	0	0.18 (0.25)	0.15 (0.40)	0.15 (0.35)	0.12 (0.17)	0.10 (0.34)	0.10 (0.17)	0.10 (0.14)	0.08 (0.17)	0.08 (0.13)
		1	0.19 (0.28)	0.19 (0.39)	0.16 (0.34)	0.12 (0.18)	0.11 (0.32)	0.11 (0.23)	0.10 (0.14)	0.09 (0.30)	0.09 (0.14)
		2	0.20 (0.30)	0.24 (0.38)	0.20 (0.30)	0.13 (0.20)	0.13 (0.27)	0.11 (0.27)	0.10 (0.15)	0.09 (0.26)	0.09 (0.22)
	4	0	0.18 (0.20)	0.16 (0.33)	0.16 (0.19)	0.12 (0.13)	0.11 (0.13)	0.11 (0.13)	0.09 (0.11)	0.09 (0.11)	0.09 (0.11)
		2	0.19 (0.22)	0.19 (0.35)	0.18 (0.24)	0.12 (0.14)	0.12 (0.17)	0.12 (0.14)	0.10 (0.12)	0.09 (0.12)	0.09 (0.11)
		4	0.20 (0.24)	0.27 (0.32)	0.21 (0.31)	0.13 (0.16)	0.13 (0.29)	0.12 (0.17)	0.10 (0.13)	0.10 (0.22)	0.10 (0.13)
	6	0	0.17 (0.18)	0.15 (0.16)	0.16 (0.17)	0.11 (0.12)	0.10 (0.11)	0.10 (0.11)	0.09 (0.09)	0.08 (0.09)	0.08 (0.09)
		3	0.18 (0.19)	0.18 (0.22)	0.17 (0.19)	0.12 (0.12)	0.12 (0.13)	0.11 (0.12)	0.09 (0.10)	0.09 (0.10)	0.09 (0.10)
		6	0.19 (0.21)	0.27 (0.30)	0.20 (0.26)	0.13 (0.14)	0.13 (0.20)	0.13 (0.14)	0.10 (0.11)	0.10 (0.11)	0.10 (0.11)
12	2	0	0.23 (0.37)	0.23 (0.54)	0.17 (0.42)	0.14 (0.23)	0.12 (0.41)	0.12 (0.34)	0.11 (0.19)	0.10 (0.36)	0.10 (0.26)
		1	0.24 (0.37)	0.27 (0.53)	0.20 (0.41)	0.15 (0.25)	0.14 (0.41)	0.12 (0.31)	0.11 (0.20)	0.10 (0.33)	0.10 (0.30)
		2	0.25 (0.34)	0.27 (0.49)	0.23 (0.40)	0.15 (0.28)	0.19 (0.42)	0.13 (0.28)	0.12 (0.22)	0.12 (0.29)	0.10 (0.26)
	4	0	0.23 (0.29)	0.25 (0.42)	0.19 (0.39)	0.15 (0.19)	0.13 (0.37)	0.13 (0.19)	0.12 (0.15)	0.11 (0.25)	0.11 (0.15)
		2	0.24 (0.32)	0.30 (0.42)	0.22 (0.37)	0.15 (0.20)	0.16 (0.34)	0.14 (0.28)	0.12 (0.16)	0.12 (0.33)	0.11 (0.16)
		4	0.26 (0.34)	0.31 (0.46)	0.28 (0.34)	0.16 (0.22)	0.23 (0.30)	0.15 (0.29)	0.12 (0.18)	0.13 (0.28)	0.12 (0.28)
	6	0	0.23 (0.25)	0.23 (0.40)	0.20 (0.29)	0.15 (0.17)	0.13 (0.20)	0.13 (0.16)	0.12 (0.13)	0.11 (0.13)	0.11 (0.13)
		3	0.24 (0.28)	0.30 (0.38)	0.23 (0.36)	0.15 (0.18)	0.16 (0.33)	0.15 (0.18)	0.12 (0.14)	0.12 (0.17)	0.12 (0.14)
		6	0.26 (0.31)	0.32 (0.35)	0.30 (0.34)	0.16 (0.20)	0.24 (0.31)	0.16 (0.29)	0.13 (0.16)	0.13 (0.30)	0.12 (0.16)
16	2	0	0.27 (0.41)	0.32 (0.53)	0.20 (0.52)	0.16 (0.32)	0.14 (0.50)	0.13 (0.38)	0.13 (0.23)	0.10 (0.43)	0.10 (0.34)
		1	0.28 (0.40)	0.31 (0.50)	0.23 (0.50)	0.17 (0.32)	0.16 (0.49)	0.14 (0.36)	0.13 (0.26)	0.12 (0.43)	0.11 (0.31)
		2	0.29 (0.39)	0.30 (0.45)	0.24 (0.45)	0.17 (0.29)	0.20 (0.45)	0.16 (0.34)	0.13 (0.26)	0.17 (0.44)	0.11 (0.27)
	4	0	0.28 (0.39)	0.38 (0.58)	0.23 (0.42)	0.17 (0.24)	0.15 (0.40)	0.15 (0.36)	0.13 (0.19)	0.12 (0.36)	0.12 (0.21)
		2	0.29 (0.40)	0.38 (0.59)	0.27 (0.41)	0.18 (0.26)	0.18 (0.40)	0.16 (0.34)	0.14 (0.21)	0.14 (0.33)	0.13 (0.32)
		4	0.31 (0.37)	0.37 (0.55)	0.29 (0.40)	0.18 (0.30)	0.24 (0.43)	0.18 (0.30)	0.14 (0.23)	0.20 (0.29)	0.13 (0.28)
	6	0	0.28 (0.33)	0.39 (0.44)	0.24 (0.41)	0.17 (0.22)	0.16 (0.38)	0.15 (0.21)	0.14 (0.17)	0.12 (0.37)	0.12 (0.17)
		3	0.30 (0.37)	0.40 (0.48)	0.28 (0.39)	0.18 (0.23)	0.19 (0.35)	0.17 (0.34)	0.14 (0.18)	0.15 (0.34)	0.13 (0.19)
		6	0.32 (0.37)	0.39 (0.56)	0.32 (0.36)	0.19 (0.26)	0.26 (0.32)	0.19 (0.31)	0.15 (0.20)	0.21 (0.30)	0.14 (0.29)

Notes: Frobenius distance between the VAR coefficients and their estimates relative to the Frobenius norm of the true coefficients for the VECIM and, in parentheses, for the VECM.

of the traditional IC. The results, reported in Table 6, show that all the criteria favor the VECIM, having values of each IC systematically lower than the VECM.

Concentrating on the outcome of the best performing IC in our Monte Carlo study, the HQIC selects a VECIM that is not only more parsimonious than the VECM, but provides evidence of a unique stochastic trend in the common components of the series. Instead, the VECM specification suggests the existence of eight permanent shocks, which is at odds with the common view that few stochastic trends lead the economy in the long run.

4.2 Contribution of shocks to the US business cycle

We rely on HQIC for model selection of the VECIM and set $p = 2$, $r = 4$ and $q = 5$. As shown in Section 2.1, we can exploit our model's features and decompose variables into two components, one of which is common, χ_t, and the other uncommon, ι_t. Afterwards, we can disentangle the part of the common component which has permanent effects, π_t, from another that is instead transitory, τ_t. Furthermore, being π_t, τ_t, and ι_t not cross-correlated at any lead and lag, we can measure the contribution of each component to the variability

Table 4: AMSFE of VECIM (VECM)

			T =240			T = 480			T =720		
N	q	r	AIC	BIC	HQIC	AIC	BIC	HQIC	AIC	BIC	HQIC
8	2	0	1.45 (1.97)	1.03 (1.20)	1.03 (1.61)	1.14 (1.33)	1.03 (1.11)	1.02 (1.07)	1.11 (1.20)	1.05 (1.08)	1.05 (1.07)
		1	1.40 (1.94)	1.07 (1.23)	1.07 (1.54)	1.11 (1.31)	1.05 (1.14)	1.04 (1.15)	1.09 (1.19)	1.05 (1.14)	1.05 (1.08)
		2	1.29 (1.75)	1.11 (1.30)	1.09 (1.43)	1.07 (1.23)	1.05 (1.12)	1.03 (1.12)	1.09 (1.18)	1.07 (1.14)	1.07 (1.12)
	4	0	1.82 (2.10)	1.03 (1.40)	1.06 (1.34)	1.22 (1.30)	1.01 (1.02)	1.01 (1.05)	1.05 (1.12)	0.99 (1.00)	0.99 (1.00)
		2	1.47 (1.77)	1.10 (1.42)	1.13 (1.41)	1.12 (1.22)	1.05 (1.08)	1.04 (1.10)	1.02 (1.07)	0.99 (1.00)	0.99 (1.01)
		4	1.26 (1.54)	1.21 (1.42)	1.14 (1.48)	1.06 (1.14)	1.04 (1.17)	1.03 (1.10)	1.03 (1.07)	1.02 (1.07)	1.02 (1.04)
	6	0	2.29 (2.24)	1.06 (1.07)	1.18 (1.37)	1.29 (1.30)	1.07 (1.08)	1.08 (1.09)	1.12 (1.14)	1.01 (1.01)	1.01 (1.02)
		3	1.50 (1.68)	1.15 (1.24)	1.26 (1.38)	1.17 (1.22)	1.09 (1.10)	1.10 (1.13)	1.03 (1.06)	1.00 (1.01)	1.00 (1.02)
		6	1.18 (1.27)	1.22 (1.31)	1.17 (1.30)	1.03 (1.06)	1.04 (1.11)	1.03 (1.05)	1.04 (1.06)	1.04 (1.06)	1.04 (1.05)
12	2	0	1.84 (2.64)	1.05 (1.56)	1.04 (1.82)	1.20 (1.44)	1.02 (1.16)	1.02 (1.24)	1.10 (1.22)	1.01 (1.09)	1.01 (1.09)
		1	1.59 (2.48)	1.05 (1.50)	1.03 (1.64)	1.20 (1.47)	1.07 (1.24)	1.07 (1.23)	1.09 (1.23)	1.04 (1.11)	1.04 (1.12)
		2	1.62 (2.32)	1.10 (1.28)	1.11 (1.46)	1.15 (1.45)	1.08 (1.27)	1.06 (1.22)	1.05 (1.20)	1.02 (1.10)	1.02 (1.08)
	4	0	2.17 (2.72)	1.08 (1.29)	1.05 (2.69)	1.23 (1.45)	1.01 (1.14)	1.01 (1.07)	1.11 (1.21)	1.02 (1.09)	1.02 (1.04)
		2	1.78 (2.67)	1.17 (1.43)	1.12 (2.16)	1.18 (1.44)	1.05 (1.24)	1.05 (1.30)	1.11 (1.21)	1.06 (1.17)	1.05 (1.10)
		4	1.55 (2.30)	1.25 (1.53)	1.22 (1.79)	1.13 (1.35)	1.11 (1.24)	1.08 (1.25)	1.05 (1.18)	1.04 (1.13)	1.03 (1.13)
	6	0	2.57 (2.65)	1.06 (1.53)	1.06 (2.15)	1.37 (1.50)	1.04 (1.09)	1.05 (1.09)	1.15 (1.21)	1.00 (1.02)	1.00 (1.02)
		3	1.97 (2.52)	1.25 (1.60)	1.22 (2.30)	1.22 (1.42)	1.10 (1.32)	1.10 (1.20)	1.08 (1.17)	1.04 (1.08)	1.03 (1.08)
		6	1.39 (1.92)	1.29 (1.49)	1.27 (1.73)	1.10 (1.26)	1.14 (1.28)	1.08 (1.28)	1.07 (1.15)	1.06 (1.16)	1.05 (1.11)
16	2	0	2.20 (3.12)	1.10 (2.00)	1.06 (2.72)	1.23 (1.65)	0.98 (1.35)	0.98 (1.20)	1.09 (1.27)	0.98 (1.11)	0.98 (1.10)
		1	2.01 (3.04)	1.11 (1.48)	1.10 (2.21)	1.23 (1.66)	1.02 (1.36)	1.02 (1.20)	1.06 (1.28)	0.98 (1.15)	0.98 (1.07)
		2	2.01 (2.96)	1.12 (1.22)	1.14 (1.44)	1.14 (1.52)	1.02 (1.22)	1.01 (1.18)	1.04 (1.24)	1.00 (1.17)	0.99 (1.06)
	4	0	2.62 (3.35)	1.16 (2.46)	1.07 (2.88)	1.39 (1.68)	1.07 (1.20)	1.07 (1.53)	1.17 (1.32)	1.04 (1.13)	1.04 (1.11)
		2	2.33 (3.66)	1.24 (2.30)	1.17 (2.65)	1.27 (1.64)	1.06 (1.28)	1.06 (1.38)	1.13 (1.28)	1.05 (1.17)	1.04 (1.21)
		4	2.03 (3.06)	1.22 (1.51)	1.31 (2.10)	1.19 (1.55)	1.10 (1.33)	1.08 (1.30)	1.10 (1.28)	1.09 (1.18)	1.05 (1.18)
	6	0	3.36 (3.55)	1.18 (1.45)	1.07 (3.68)	1.45 (1.64)	1.04 (1.17)	1.04 (1.13)	1.21 (1.30)	1.02 (1.14)	1.02 (1.04)
		3	2.50 (3.73)	1.32 (1.89)	1.26 (2.84)	1.31 (1.63)	1.07 (1.30)	1.09 (1.52)	1.11 (1.26)	1.04 (1.18)	1.04 (1.10)
		6	2.03 (2.94)	1.35 (1.75)	1.42 (2.28)	1.18 (1.50)	1.15 (1.34)	1.11 (1.40)	1.11 (1.25)	1.11 (1.20)	1.07 (1.21)

Notes: Average of the mean square 1−step ahead forecast errors over the n series for the VECIM and, in parentheses, for the VECM.

Table 5: Percentages of the correct estimation of q for the VECIM when $q = n$

			T =240			T = 480			T = 720		
N	q	r	AIC	BIC	HQIC	AIC	BIC	HQIC	AIC	BIC	HQIC
12	12	0	100.0	99.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		3	100.0	97.4	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		6	100.0	95.2	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Percentages with which each IC correctly estimates the true number of indexes q when $q = n$

of the k−th element of Y_t at a specific frequency band. In order to evaluate the sample variability of our estimates, we implement a bootstrap procedure by sampling 2000 times with replacement from the VECIM residuals keeping the specification of the model fixed over replications.

Table 7 shows the percentage of the variance of each variable that is explained by each component at the business cycle frequency band, i.e. $\lambda \in [2\pi/32, 2\pi/6]$. The common component explains most of the cyclical

Table 6: Information Criteria for the VECIM and the VECM

	AIC	BIC	HQIC
VECIM	-17.351	-15.984	-16.586
p, r, q	4,5,6	3,0,1	2,4,5
VECM	-17.254	-15.773	-16.296
p, r	2,6	1,2	2,2

Notes: p, r, q [p, r] indicate the selected specification of the VECIM [VECM] for each IC.

variability of all the variables taken into account. The shares of the uncommon component are less than 20% for all variables but Labor Productivity and TFP, which are, respectively, around 35% and 59%.

Table 7: Variance contribution of each component at the business cycle frequencies

	Common	Permanent	Transitory
Unemployment	97.5 [95.6, 98.7]	23.9 [7.8, 40.5]	76.1 [59.5, 92.2]
Output	95.6 [92.6, 97.6]	31.2 [13.3, 48.5]	68.8 [51.5, 86.7]
Hours Worked	86.7 [80.4, 91.4]	25.4 [8.2, 42.6]	74.6 [57.4, 91.8]
Investment	95.8 [93.1, 97.6]	24.2 [7.8, 42.4]	75.8 [57.6, 92.2]
Consumption	94.5 [87.8, 98.0]	34.3 [21.5, 48.1]	65.7 [51.9, 78.5]
TFP	41.2 [28.2, 54.2]	10.6 [3.1, 27.0]	89.4 [73.0, 96.9]
Labor Prod.	65.2 [53.7, 75.8]	25.8 [9.7, 46.1]	74.2 [53.9, 90.3]
Labor Share	92.9 [85.3, 97.0]	11.8 [3.8, 26.2]	88.2 [73.8, 96.2]
Inflation	80.2 [68.5, 89.5]	33.5 [11.8, 57.0]	66.5 [43.0, 88.2]
Nom. Int. rate	87.1 [79.5, 92.5]	21.0 [5.8, 42.5]	79.0 [57.5, 94.2]

Notes: Medians and, in brackets, the 16th and 84th percentiles of the variance contribution at the business cycle frequencies of each component. The distributions are obtained with bootstrap.

Within the common component the transitory shocks play a more important role than the permanent one, but the latter still explains a non-negligible portion of variation in the short run with values ranging between 11% and 34%. Surprisingly, the highest shares of the transitory component are for TFP and Labour Share, with more than 88% of explained cyclical variability.

4.3 Identification of the business cycle driver

The source of the business cycle is of great interest in macroeconomics. Understanding what drives macroeconomic fluctuations in the short-run is a key information for policy makers. The standard approach suggests to identify structural shocks consists in imposing some restrictions coming from economic theory on the reduced form errors. Afterwords, the assessment of the features of each structural shock is done by computing its impulse response function (IRF) and forecasting error variance decomposition at various time horizons. Clearly, the validity of such procedure strongly depends on the underlying restrictions. The fact that the results may largely vary depending on *ex-ante* economic assumptions has led a branch of the literature to propose an "agnostic" methodology, based on the maximization of the variance of a given variable in a set of time horizons or in a frequency band. The max-share identification strategy has been originally proposed by Uhlig (2004) and further extended by other authors, among which Barsky and Sims (2011), Francis et al. (2014) and Angeletos et al. (2020). In particular, the latter authors identify the Main Business Cycle shock (MBC) as the one that maximes the variability of the the unemployment rate at the business cycle frequencies, and claim that it behaves as neither a standard technology shock nor a classic demand type shock, because it displays a disconnection to both TFP and inflation. For this reason, they interpret it as a demand shock not affecting nominal rigidities.[12]

Our empirical application builds on Angeletos et al. (2020) and aims to give additional insight on the transmission of such shock. For structural analysis, we proceed as follows. Let us first focus on the spectral density matrix of the structural representation of process Y_t

$$F(\lambda) = \frac{1}{2\pi}\tilde{\Psi}^*(z)HH'\tilde{\Psi}^*(z^{-1})', \tag{19}$$

where $\tilde{\Psi}^*(z) = \Psi^*(z)S$, $SS' = \Omega$, H is an $n \times n$ matrix such that $HH' = I$ and the structural shocks are $H'S^{-1}\varepsilon_t$. Hence, the contribution of the generic $j-$th structural shock to the spectral density of the $k-$th element of series Y_t at frequency λ is

$$f(k, j, \lambda) = \frac{1}{2\pi}\tilde{\psi}_k^*(z)'hh'\tilde{\psi}_k^*(z^{-1}) = \frac{1}{2\pi}h'\tilde{\psi}_k^*(z^{-1})\tilde{\psi}_k^*(z)'h \equiv h'\Theta(k, \lambda)h, \tag{20}$$

where $\tilde{\psi}_k^*(z)'$ is the $k-$th row of the matrix $\tilde{\Psi}^*(z)$, and h is the $j-$th column of H.[13]

Angeletos et al. (2020) propose to identify the $j-$th structural shock in an atheoretical way, ie, as the shock that maximizes the variability of $k-$th series at the frequency band $[\lambda_a, \lambda_b]$, with $0 < \lambda_a \leq \lambda_b \leq \pi$. In view of Equation (20), this requires one to choose h as the eigenvector corresponding to the largest eigenvalue of the matrix

$$\Theta(k, \lambda_a, \lambda_b) = \frac{1}{2\pi}\int_{\lambda_a}^{\lambda_b} \tilde{\psi}_k^*(z^{-1})\tilde{\psi}_k^*(z)'d\lambda$$

When $[\lambda_a, \lambda_b] = [2\pi/32, 2\pi/6]$ and $k-$th series is a procyclical variable, Angeletos et al. (2020) label the structural shock $h'S^{-1}\varepsilon_t$ as the MBC and its IRF is obtained as $\tilde{\Psi}^*(L)h$.

Similarly, we can apply the same procedure on the common component of the $k-$th variable and obtain the Main Common Business Cycle shock (MCBC). This is motivated by the fact that, if the model correctly identifies the common errors as those that span the structural shocks, the uncommon errors should basically be noise that do not contain additional economic information, see *e.g.* Cubadda and Hecq (2022a).

Specifically, the spectral density matrix of the structural representation of the common component χ_t is

$$F_\chi(\lambda) = \frac{1}{2\pi}\tilde{C}^*(z)H_\chi'H_\chi\tilde{C}^*(z^{-1})',$$

where $\tilde{C}^*(z) = C^*(z)S_\chi$, $C^*(z) = \Psi^*(z)\Omega\omega\Sigma^{-1}$, $S_\chi S_\chi' = \Sigma$, and H_χ is the $q \times q$ orthonormal matrix that identifies the structural common shock. The $j-$th column of the matrix H_χ, denoted as h_χ, is the eigenvector corresponding to the largest eigenvalue of the matrix

$$\Theta_\chi(k, 2\pi/32, 2\pi/6) = \frac{1}{2\pi}\int_{2\pi/32}^{2\pi/6} \tilde{c}_k^*(z^{-1})\tilde{c}_k^*(z)'d\lambda,$$

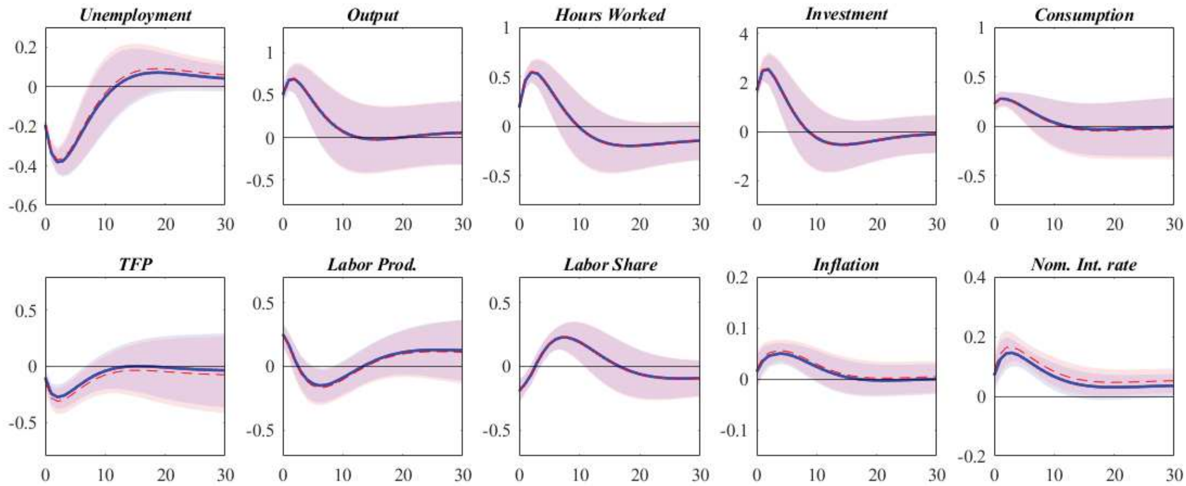
where $\tilde{c}_k^*(z)'$ is the $k-$th row of the matrix $\tilde{C}^*(z)$. Hence, the MCBC is $h_\chi'S_\chi^{-1}\omega'\varepsilon_t$ and its IRF is $\tilde{C}^*(L)h_\chi$.

[12] Angeletos et al. (2020) show that alternatively targeting unemployment, output, hours worked, consumption and investment, their approach provides very similar impulse responses functions for the variables of interest.

[13] The orthogonality of the structural shocks implies that the spectral density of the $k-$th series is equal to $\sum_{j=1}^n f(k, j, \lambda)$.

Following Angeletos et al. (2020), we use the unemployment rate as the target variable, i.e. the $k-$th series for which the variability [of its common component] at the business cycle frequencies is maximized by the MBC [MCBC]. Figure (1) compares the IRFs that are respectively associated with the MBC and the MCBC. The sign of such shocks is established so to have a positive effect on GDP on impact.

Figure 1: Impulse responses of MBC and MCBC



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 1 standard deviation confidence bands obtained with bootstrap.

Overall, the responses of the MBC and the MCBC are almost indistinguishable, confirming that the uncommon errors are basically noise that contain no relevant information for structural analysis. Moreover, we observe a pro-cyclical response to the M(C)BC of investment, output, consumption, hours worked, and unemployment. Nevertheless, the responses of both inflation and TFP are quite flat, thus apparently precluding a straightforward interpretation as either a traditional demand shock or a supply shock.

Table (8) shows the portion of the business cycle variability explained by the MBC and MCBC of, respectively, the variables and their common components. The MBC explains most of the cyclical variation of almost all the variables, whereas leave unexplained some of them, in particular inflation, TFP and labor productivity. When considering the MCBC, we see that these conclusions are essentially confirmed for output, unemployment, consumption, investment, hours worked and labor share, whereas the picture changes for labor productivity, and, especially, TFP, whose relative explained cyclical variability improves from 17% to 48%. Recalling that TFP and Labor Productivity are characterized by a smaller common component w.r.t. the other variables, the superior share of the MCBC is due to the exclusion of the uncommon component, which de-emphasizes the role of the MBC on TFP and apparently causes the disconnection of such variable from the business cycle documented by Angeletos et al. (2020). Regarding the disconnection of inflation, an analogous reasoning does not apply since the MCBC and MBC display similar shares, although we observe a larger contribution of such shocks to the short-run variability of inflation w.r.t. Angeletos et al. (2020)

Let us complete the portrait of the identified shocks by looking at columns 3 and 4 of Table (8), which reports the variance contributions at the long run frequency band. Since the spectral density of I(1) variables is unbounded at the zero frequency, we consider the frequencies corresponding to quarters in the range

Table 8: Variance contribution of MBC and MCBC

	Short-run		Long-run	
	MBC	MCBC	MBC	MCBC
Unemployment	66.1 [56.1, 78.3]	65.9 [55.9, 78.3]	24.7 [6.9, 63.5]	30.9 [9.3, 74.5]
Output	55.0 [44.5, 66.9]	59.5 [48.0, 72.4]	12.3 [2.4, 47.4]	14.9 [2.9, 56.7]
Hours Worked	50.2 [40.0, 61.0]	60.0 [48.5, 72.0]	17.0 [5.9, 37.6]	29.5 [9.7, 64.3]
Investment	58.7 [47.2, 70.2]	63.4 [50.4, 76.0]	14.2 [3.9, 41.7]	19.2 [5.0, 56.0]
Consumption	28.6 [15.0, 45.9]	31.3 [16.4, 49.9]	9.0 [1.2, 38.7]	10.4 [1.3, 43.5]
TFP	17.0 [6.8, 29.5]	48.1 [22.4, 70.1]	5.6 [0.5, 28.3]	12.6 [1.1, 54.0]
Labor Prod.	22.8 [13.3, 34.3]	36.7 [22.9, 50.8]	6.5 [0.6, 31.2]	11.9 [1.0, 51.9]
Labor Share	30.9 [18.7, 44.4]	34.7 [21.6, 48.8]	13.5 [3.3, 37.5]	20.0 [4.7, 53.9]
Inflation	13.4 [4.9, 26.7]	18.8 [7.5, 35.5]	6.8 [1.4, 21.0]	18.2 [3.9, 50.9]
Nom. Int. rate	33.9 [16.9, 58.8]	46.3 [24.2, 73.7]	13.6 [4.7, 30.4]	33.5 [11.1, 65.4]

Notes: Medians and, in brackets, the 16th and 84th percentiles of the variance contribution of the MBC and the MCBC at the various frequency bands. The distributions are obtained with bootstrap.

$[80, 256]$. The contribution of the MBC at these frequencies is larger than what was found by Angeletos et al. (2020) but still small. As for the short run, the MCBC explains relatively more variability than the MBC also in longer periods.

A possible reason why the MCBC has not a straightforward economic interpretation may lie in the fact that the atheoretical identification scheme picks up a linear combination of two structural shocks with a different nature, e.g. a demand shock and productivity shock, both of them having a non negligible role in determining the business cycle fluctuations (Dieppe et al., 2021). Hence, we aim at identifying the Main Common Transitory Business Cycle shock (MCTBC) as the shock that maximizes the cyclical variability of the common-transitory component of unemployment. Formally, the spectral density matrix of the structural representation of the common transitory component τ_t is

$$F_\tau(\lambda) = \frac{1}{2\pi}\tilde{T}^*(z)H'_\tau H_\tau \tilde{T}^*(z^{-1})',$$

where $\tilde{T}^*(z) = C^*(z)\underline{\alpha}_0(\underline{\alpha}'_0\Sigma^{-1}\underline{\alpha}_0)^{-1}S_\chi$, $S_\chi S'_\chi = \underline{\alpha}'_0\Sigma^{-1}\underline{\alpha}_0$, and H_τ is the $r \times r$ orthonorrmal matrix that identifies the structural common transitory shocks. The $j-$th column of the matrix H_τ, denoted as h_τ, is then the eigenvector corresponding to the largest eigenvalue of the matrix

$$\Theta_\tau(k, 2\pi/32, 2\pi/6) = \frac{1}{2\pi}\int_{2\pi/32}^{2\pi/6} \tilde{t}^*_k(z^{-1})\tilde{t}^*_k(z)'d\lambda,$$

where $\tilde{t}^*_k(z)'$ is the $k-$th row of the matrix $\tilde{T}^*(z)$. Hence, the MCTBC is $h'_\tau S_\tau^{-1}\underline{\alpha}'_0\Sigma^{-1}\omega'\varepsilon_t$ and its IRF is $\tilde{T}^*(L)h_\chi$.

An analogous strategy can be used to identify the main common permanent business cycle shock but, given that in our empirical model we set $q = 5$ and $r = 4$, we have a unique common permanent shock. Thus, the Main Common Permanent Business Cycle shock (MCPBC) is simply proportional to such shock.

As we can observe in Figure 2, the newly identified shocks have a much more insightful behavior than the M(C)BC. In particular, the IRFs of the MCTBC confirm its temporary nature since, after a peak in the first period, the effects on variables rapidly decay towards zero. As for MBC and MCBC, inflation and nominal interest rate react positively to MCTBC, even though the effect on the former is not particularly large, whereas the impact on TFP and labor productivity is negative for the horizons where it is significant. Instead, the MCPBC triggers a long-lasting effect on most of the variables, especially GDP, consumption, TFP and productivity. Differently from its transitory counterpart, the impact of MCPBC on inflation and nominal interest rate is significantly negative at shorter horizons and then decays towards zero after some periods. Remarkably, the contemporaneous impact on TFP is small and not significant and it tends to steadily increase after 3 quarters. All in all, the MCTBC exhibits the main features of a demand shock whereas the MCPBC may be interpreted as a news shock à la Beaudry and Portier (2014).

The contributions over the common component variability are displayed in Table 9. We see that the MCTBC explains a sizeable portion of the common component variation in the short run, whereas only a small part in the long run, in particular for TFP, labor productivity, GDP and consumption. Analogously, the MCPBC covers only a modest, but still non-negligible, part of the common component cyclical volatility. Consumption, output and inflation display the highest figures in this setting, with a share of cyclical variability explained by the MCPBC that is larger than 31%. Remarkably, consumption is explained more by MCPBC than by MCTBC in short horizons; a result that is consistent with permanent income theories, according which volatility in consumption is mostly due to fluctuations in permanent income. Furthermore, the MCPBC plays a clear dominant role in the long-run frequency band, confirming that we have successfully identified a permanent and a transitory shock.

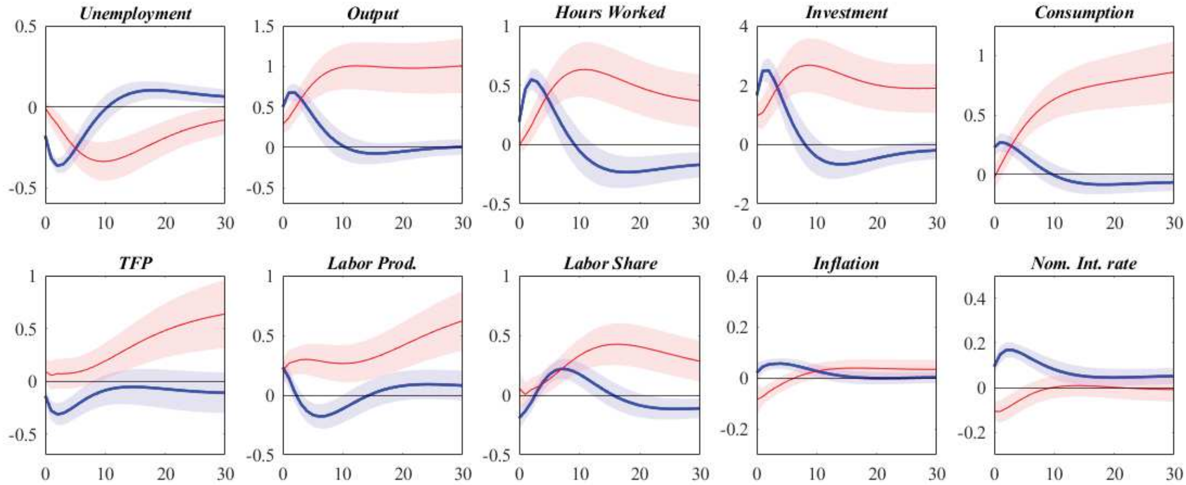
Table 9: Variance contribution of MCPBC and MCTBC

	Short-run		Long-run	
	MCPBC	MCTBC	MCPBC	MCTBC
Unemployment	23.9 [7.8, 40.5]	59.9 [46.5, 73.2]	56.2 [18.6, 83.2]	15.8 [4.8, 48.4]
Output	31.2 [13.3, 48.5]	51.4 [35.9, 65.7]	87.6 [54.1, 95.9]	3.4 [0.9, 13.7]
Hours Worked	25.4 [8.2, 42.6]	53.8 [40.1, 67.4]	55.2 [16.4, 84.3]	18.4 [5.7, 46.1]
Investment	24.2 [7.8, 42.4]	55.5 [40.7, 69.6]	68.6 [24.3, 89.7]	8.8 [2.6, 27.3]
Consumption	34.3 [21.5, 48.1]	27.6 [11.7, 43.5]	91.4 [75.3, 95.9]	2.4 [0.6, 9.3]
TFP	10.6 [3.1, 27.0]	48.0 [23.4, 70.2]	45.3 [4.9, 80.8]	4.4 [0.5, 25.2]
Labor Prod.	25.8 [9.7, 46.1]	33.1 [19.9, 47.0]	67.6 [24.0, 87.9]	3.4 [0.5, 18.4]
Labor Share	11.8 [3.8, 26.2]	29.2 [16.6, 42.2]	54.6 [13.4, 84.6]	10.4 [2.4, 32.4]
Inflation	33.5 [11.8, 57.0]	19.7 [8.0, 34.6]	21.5 [3.5, 62.6]	12.6 [2.8, 38.7]
Nom. Int. rate	21.0 [5.8, 42.5]	49.3 [30.4, 67.9]	14.7 [1.9, 49.3]	29.8 [9.1, 63.3]

Notes: Medians and, in brackets, the 16th and 84th percentiles of the variance contribution of the MCPBC and the MCTBC at the various frequency bands. The distributions are obtained with bootstrap.

Interestingly, our findings are to some extent in agreement with those of Avarucci et al. (2022), who, using a DFM approach, document that the bulk of the business cyle fluctuations is explained by a transitory shock, and, to a minor extent, by a permanent shock. On the basis of variance decomposition at different frequency bands, they argue that it is tempting to interpret the former as a demand shock and the latter as a supply shock. However, Avarucci et al. (2022) do not provide estimates of the impulse response function of those shocks, in absence of which it is somewhat cumbersome to corroborate economic interpretations. Moreover, the variability at frequency zero of series ΔY_t that is explained by the shock that Avarucci et al. (2022) label as transitory is relatively small but not null by construction as in the case of the MCTBC.

Figure 2: Impulse responses of MCPBC and MCTBC



Notes: the red solid line is the point estimation of the responses to the MCPBC, the blue solid line is the point estimation of the responses to the MCTBC whereas the shaded areas are the 1 standard deviation confidence bands obtained with bootstrap.

5 Conclusions

Modern time series econometrics has stressed the advantages of departing from the traditional small-scale models for both forecasting and structural analysis. Unfortunately, standard inference for cointegrated VARs becomes unreliable as we include more variables in the analysis. In order to achieve dimension reduction for cointegrated VARs of medium dimensions, we propose a novel model consisting in a particular reduced-rank structure on the parameters of the VECM. The resulting formulation, i.e. the VECIM, allows to parsimoniously express the first differences of a set cointegrated time series as a linear function of the lags of a smaller number of observable factors. Moreover, the VECIM enables to decompose variables into three components that are uncorrelated at any lag or lead: a common permanent component, a common transitory component, and an uncommon component. Hence, the new model combines an attractive feature of the DFM, i.e. disentangling the shocks that are common among variables, with one of the VECM, i.e. decomposing shocks into permanent and transitory ones.

We provide a switching algorithm for Gaussian ML estimation of the VECIM parameters, we opt for informational methods to specify the model, and we assess the finite sample performances of the proposed methodology by means of a comprehensive Monte Carlo study. The results show that our model outperforms the VECM when an index structure is present in the data. This outcome becomes clearer as the number of variables increases.

Finally, we use the peculiar features of our model in order to identify the shocks that drive the business cycle. On the on hand, we decompose a set of key US macro variables in all the aforementioned components, on the other hand, we identify two shocks: one that explains the bulk of variation of the common transitory component of unemployment at the business cycle frequencies, and the second one which does the same but for the common permanent component of unemployment. These two shocks reveal to be endowed with a neater economic interpretation than a unique main business cycle shock identified according to Angeletos et al. (2020).

Overall, the method is not difficult to implement and promises to be useful in both handling medium datasets and adding insights to structural analysis.

References

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

Avarucci, M., Cavicchioli, M., Forni, M., and Zaffaroni, P. (2022). The main business cycle shock(s): Frequency-band estimation of the number of dynamic factors. *CEPR Press Discussion Paper*, (17281).

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, 122(1):137–183.

Bai, J. and Ng, S. (2004). A PANIC attack on unit roots and cointegration. *Econometrica*, 72(4):1127–1177.

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

Barigozzi, M., Lippi, M., and Luciani, M. (2021). Large-dimensional dynamic factor models: Estimation of impulse–response functions with I(1) cointegrated factors. *Journal of Econometrics*, 221(2):455–482.

Barigozzi, M. and Trapani, L. (2022). Testing for common trends in nonstationary large datasets. *Journal of Business & Economic Statistics*, 40(3):1107–1122.

Barsky, R. B. and Sims, E. R. (2011). News shocks and business cycles. *Journal of Monetary Economics*, 58(3):273–289.

Beaudry, P. and Portier, F. (2014). News-driven business cycles: Insights and challenges. *Journal of Economic Literature*, 52(4):993–1074.

Bernardini, E. and Cubadda, G. (2015). Macroeconomic forecasting and structural analysis through regularized reduced-rank regression. *International Journal of Forecasting*, 31(3):682–691.

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*, 37,(4):700–721.

Carriero, A., Kapetanios, G., and Marcellino, M. (2011). Forecasting large datasets with Bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, 26(5):735–761.

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

Casoli, C. and Lucchetti, R. (2022). Permanent-transitory decomposition of cointegrated time series via dynamic factor models, with an application to commodity prices. *Econometrics Journal*, 25(2):494–514.

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.

Cavaliere, G., Rahbek, A., and Taylor, A. (1996). Bootstrap determination of the cointegration rank in vector autoregressive models. *Econometrica*, 80(4):1721–1740.

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.

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. (2011). Testing for common autocorrelation in data-rich environments. *Journal of Forecasting*, 30(3):325–335.

Cubadda, G. and Hecq, A. (2022a). Dimension reduction for high dimensional vector autoregressive models. *Oxford Bulletin of Economics and Statistics*, 84(5):1123–1152.

Cubadda, G. and Hecq, A. (2022b). Reduced rank regression models in economics and finance. *Oxford Research Encyclopedia of Economics and Finance.*

Cubadda, G., Hecq, A., and Palm, F. (2009). Studying co-movements in large multivariate data prior to multivariate modelling. *Journal of Econometrics*, 148(1):25–35.

Dieppe, A., Francis, N., and Kindberg-Hanlon, G. (2021). The identification of dominant macroeconomic drivers: Coping with confounding shocks. *ECB Working Paper*, (2534).

Diniz, M., Pereira, C., and Stern, J. (2020). Cointegration and unit root tests: A fully bayesian approach. *Entropy*, 22(9):1–23.

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., Lippi, M., and Sala, L. (2020). Common component structural VARs. *CEPR Press Discussion Paper*, (15529).

Forni, M., Gambetti, L., and Sala, L. (2019). Structural VARs and noninvertible macroeconomic models. *Journal of Applied Econometrics*, 34(2):221–246.

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.

Francis, N., Owyang, M. T., Roush, J. E., and DiCecio, R. (2014). A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks. *Review of Economics and Statistics*, 96(4):638–647.

Giannone, D., Lenza, M., and Primiceri, G. E. (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics*, 97(2):436–451.

Gonzalo, J. and Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economics Statistics*, 13(1):27–35.

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.

Granger, C. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1):121–130.

Hautsch, N., Okhrin, O., and Ristig, A. (2022). Maximum-likelihood estimation using the zig-zag algorithm. *Journal of Financial Econometrics*, forthcoming.

Hecq, A., Margaritella, L., and Smeekes, S. (2021). Granger causality testing in high-dimensional VARs: A post-double-selection procedure. *Journal of Financial Econometrics*, forthcoming.

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.

Hsu, N. J., Hung, H. L., and Chang, Y. M. (2008). Subset selection for vector autoregressive processes using lasso. *Computational Statistics & Data Analysis*, 52(7):3645—-3657.

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

Kasa, K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics*, 29(1):95–124.

Kock, A. B. and Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, 186(2):325—-344.

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

Lippi, M., Deistler, M., and Anderson, B. (2022). High-dimensional dynamic factor models: A selective survey and lines of future research. *Econometrics and Statistics*, forthcoming.

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

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

Reinsel, G. and Ahn, S. (1992). Vector autoregressive models with unit roots and reduced rank structure: Estimation, likelihood ratio test, and forecasting. *Journal of Time Series Analysis*, 13(4):353–375.

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

Sims, C. A., Stock, J., and Watson, M. (1990). Inference in linear time series models with some unit roots. *Econometrica*, 58(1):113–144.

Smeekes, S. and Wijler, E. (2020). Unit roots and cointegration. In Fuleky, P., editor, *Macroeconomic Forecasting in the Era of Big Data. Advanced Studies in Theoretical and Applied Econometric*, volume 52, pages 541–584. Springer.

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.

Uhlig, H. (2004). What moves real GNP? *Econometric Society 2004 North American Winter Meetings*, 636.

Vahid, F. and Engle, R. (1993). Common trends and common cycles. *Journal of Applied Econometrics*, 8(4):341–360.

Zhang, R., Robinson, P., and Yao, Q. (2019). Identifying cointegration by eigenanalysis. *Journal of the American Statistical Association*, 114(526):916–927.

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Detecting Common Bubbles in Multivariate Mixed Causal-noncausal Models
Gianluca Cubadda, Alain Hecq and Elisa Voisin
CEIS Research Paper, 555 February 2023

The More You Breath, The Less You Are Safe. The Effect of Air Pollution on Work Accidents
Domenico Depalo and Alessandro Palma
CEIS Research Paper, 554 February 2023

Gains from Variety: Refugee-Host Interactions in Uganda
Rama Dasi Mariani, Furio Camillo Rosati, Pasquale Scaramozzino and Marco d'Errico
CEIS Research Paper, 553 February 2023

A Reinforcement Learning Algorithm for Trading Commodities
Federico Giorgi, Stefano Herzel and Paolo Pigato
CEIS Research Paper, 552 February 2023

Debt and Financial Fragility: Italian Non-Financial Companies after the Pandemic
Bassam Fattouh, Beniamino Pisicoli and Pasquale Scaramozzino
CEIS Research Paper, 551 February 2023

On the Role of Bargaining Power in Nash-in-Nash Bargaining: When More is Less
Marc Escrihuela-Villar, Walter Ferrarese and Alberto Iozzi
CEIS Research Paper, 550 December 2022

Traumatic Experiences Adversely Affect Life Cycle Labor Market Outcomes of the Next Generation - Evidence from WWII Nazi Raids.
Vincenzo Atella, Edoardo Di Porto, Joanna Kopinska and Maarten Lindeboom
CEIS Research Paper, 549 December 2022

Inter-municipal Cooperation in Public Procurement
Giampaolo Arachi, Debora Assisi, Berardino Cesi, Michele G. Giuranno and Felice Russo
CEIS Research Paper, 548, December 2022

Testing for Endogeneity of Irregular Sampling Schemes
Aleksey Kolokolov, Giulia Livieri and Davide Pirino
CEIS Research Paper, 547, December 2022

Aggregation Trees
Riccardo Di Francesco
CEIS Research Paper, 546, December 2022