Online Appendix: Structural VAR models in the Frequency Domain

Alain Guay*<br>Florian Pelgrin ${ }^{\dagger}$


#### Abstract

This online appendix provides (1) some complementary proofs regarding the main results, (2) some additional theoretical results in the case of local misspecification, (3) the derivation of the discretized ALS estimator, (4) the identification of only one structural shock of interest, (5) some further Monte Carlo simulations, (6) an overview of the spectral density matrix and its application to SVAR models, (7) the comparison between our approach and the nonparametric correction of Christiano, Eichenbaum and Vigfussion (2006a, 2006b), (8) some applications of the C-ALS approach, and (9) the unreliability issue of a long-run identification scheme.


[^0]This supplementary material presents some elements regarding:

## 1. Further theoretical derivations

We present the proof of Proposition 4.1, the derivation of the spectral decomposition of the operator $K$, and the closed form determination of the C-ALS estimator in the case of bivariate VAR models.

## 2. Misspecification and C-ALS

We provide theoretical results about the asymptotic distribution and the overidentification test in the presence of local misspecification. Furthermore, some Monte Carlo simulations are discussed.

## 3. The derivation of a discretized ALS estimator

As explained in Section 2 of the main text, a first strategy to impose identifying restrictions in the frequency domain is to consider a discretization of the frequency band of interest. After distinguishing the real part and the imaginary part of the identifying restrictions, the estimation of the structural parameters $a(0)$ can be settled down as an asymptotic least squares problem.

## 4. Identification of only one structural shock of interest

We discuss the case in which one aims at identifying only one structural shock.

## 5. Further Monte Carlo simulations

Starting from the general definition of the data generating process (Section 5 of the main text), we provide some additional results using a $\operatorname{VAR}(2)$ specification.

## 6. The spectral density matrix and structural VAR models

We briefly discuss the concept of spectral density matrix and then apply it to the case of structural VAR models.

## 7. Comparison between C-ALS and the CEV-based approach

To circumvent the uncertainty surrounding the estimation of long-run multipliers, Christiano et al.
(2006a, 2006b) have proposed implementing a nonparametric correction of the spectral density matrix at $\omega=0$. We compare their smoothing-based approach (using a local average of the periodogram) with our procedure.

## 8. Applications of the C-ALS approach

This section provides some applications of the asymptotic least squares theory using a frequency band or equivalently a continuum of estimating equations.
9. The unreliability issue of a long-run identification scheme

This section discusses the theoretical foundations of the so-called unreliability problem in the case of a long-run identification scheme.

## 1 Further theoretical derivations

In this section, we present the proof of Proposition 3.1, the derivation of the spectral decomposition of the operator $K$, and the closed form determination of the C-ALS estimator in the case of bivariate VAR models.

### 1.1 Proof of Proposition 3.1

The estimator is given by

$$
\widehat{\alpha}_{T}=\arg \min _{\alpha \in \mathcal{A}}\left\|S_{T} g\left(a, \widehat{b}_{T}, \omega\right)\right\|
$$

where $S_{T}$ is a sequence of random bounded linear operators.

First, under Assumption A. 1 to Assumption A.5, $\widehat{\alpha}_{T} \xrightarrow{p} \alpha_{0}$ by Theorem 2.1 of Newey and McFadden (1994). Now, differentiating the objective function with respect to $\alpha$ and $\beta$ by a mean value expansion leads to:

$$
\begin{gathered}
\left\langle S_{T} \frac{\partial g}{\partial \alpha^{\prime}}\left(\widehat{\alpha}_{T}, \widehat{\beta}, \omega_{T}\right), S_{T} g\left(\widehat{\alpha}_{T}, \widehat{\beta}_{T}, \omega\right)\right\rangle=0 \\
\Longleftrightarrow\left\langle S_{T} \frac{\partial g}{\partial \alpha^{\prime}}\left(\widehat{\alpha}_{T}, \widehat{\beta}_{T}, \omega\right), S_{T}\left\{g\left(\alpha_{0}, \beta_{0}, \omega\right)+\frac{\partial g}{\partial \alpha^{\prime}}\left(\bar{a}_{T}, \widehat{\beta}_{T}, \omega\right)\left(\widehat{\alpha}_{T}-\alpha_{0}\right)+\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}, \bar{\beta}_{T}, \omega\right)\left(\widehat{\beta}_{T}-\beta_{0}\right)\right\}\right\rangle=0
\end{gathered}
$$

where $\bar{a}_{T}$ is on the line segment joining $\widehat{\alpha}_{T}$ and $\alpha_{0}, \bar{\beta}_{T}$ is on the line segment joining $\widehat{\beta}_{T}$ and $\beta_{0}$ and under Assumption A.1, $g\left(\alpha_{0}, \beta_{0}, \omega\right)=0$.

Using the linearity of the operator and $g\left(\alpha_{0}, \beta_{0}, \omega\right)=0$, we obtain:
$\widehat{\alpha}_{T}-\alpha_{0}=-\left\langle S_{T} \frac{\partial g}{\partial \alpha^{\prime}}\left(\widehat{\alpha}_{T}, \widehat{\beta}_{T}, \omega\right), S_{T} \frac{\partial g}{\partial \alpha^{\prime}}\left(\bar{a}_{T}, \widehat{\beta}_{T}, \omega\right)\right\rangle^{-1}\left\langle S_{T} \frac{\partial g}{\partial \alpha^{\prime}}\left(\widehat{\alpha}_{T}, \widehat{\beta}_{T}, \omega\right), S_{T} \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}, \bar{\beta}_{T}, \omega\right)\left(\widehat{\beta}_{T}-\beta_{0}\right)\right\rangle$.
Since $\widehat{\alpha}_{T} \xrightarrow{p} \alpha_{0}, \widehat{\beta}_{T} \xrightarrow{p} \beta_{0}$ and under the assumption that $\left\|S_{T}-S\right\| \rightarrow 0$ in probability

$$
\begin{aligned}
\sqrt{T}\left(\widehat{\alpha}_{T}-\alpha_{0}\right)= & -\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right)\right\rangle^{-1}\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), S \frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right) \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right)\right\rangle \\
& +o_{p}(1)
\end{aligned}
$$

Using Lemma 4.1 and $\left\|S_{T}-S\right\| \rightarrow 0$ in probability, one has

$$
S_{T} \frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right) \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right) \Rightarrow Y
$$

and $Y=\mathcal{N}\left(0, S K \bar{S}^{\prime}\right)$.

The asymptotic variance-covariance matrix of $\sqrt{T}\left(\widehat{\alpha}_{T}-\alpha_{0}\right)$ depends on the following expression:

$$
\begin{array}{r}
E\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right), S \frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}\right) \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right)\right\rangle\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right), S \frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}\right) \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right)\right\rangle^{\prime}
\end{array}=
$$

Using

$$
K^{*}=E\left[\overline{\frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}\right)} \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right) \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right)^{\prime} \frac{\partial g^{\prime}}{\partial \beta}\left(\alpha_{0}, \beta_{0}\right)\right]=\left[\overline{\frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}\right)} \Omega \frac{\partial g^{\prime}}{\partial \beta}\left(\alpha_{0}, \beta_{0}\right)\right]
$$

and $K=K^{*}$ (the operator $K$ is self-adjoint) imply

$$
=\int_{\underline{\omega}}^{\bar{\omega}} \frac{\partial g^{\prime}}{\partial \alpha}\left(\alpha_{0}, \beta_{0}\right) S\left(\int_{\underline{\omega}}^{\bar{\omega}} S K \overline{S\left(\frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right)\right.} d \omega\right) d \omega=\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right),\left(S K \bar{S}^{\prime}\right) S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right)\right\rangle .
$$

under the assumption that $S$ is Hermitian. Then, for

$$
\sqrt{T}\left(\widehat{\alpha}_{T}-\alpha_{0}\right)=-\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right), S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right)\right\rangle^{-1}\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right), Y\right\rangle+o_{p}(1)
$$

using the previous result,

$$
\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right), Y\right\rangle \sim \mathcal{N}\left(0,\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right),\left(S K \bar{S}^{\prime}\right) S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}\right)\right\rangle\right) .
$$

The result for the asymptotic distribution for a given sequence of random linear operators $S_{T}$ follows.

### 1.2 Derivation of the spectral decomposition of the operator $K$

Proposition 1.1. Let $\widehat{\alpha}_{T}^{1}$ denote a first-step consistent estimator of $\alpha_{0}$. A consistent estimator of the Moore-Penrose generalized inverse is defined by:

$$
\left(K_{T}^{-1} f\right)\left(\omega_{1}\right)=\sum_{i=1}^{q} \frac{1}{\lambda_{i, T}} \gamma_{i, T}\left(\omega_{1}\right)\left\langle f, \gamma_{i, T}\right\rangle
$$

where $\gamma_{i, T}\left(\omega_{1}\right)$ is given by:

$$
\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \widehat{\Omega}_{T}^{1 / 2} D_{i, T}
$$

the eigenvalues $\lambda_{i, T}$ are those of the $q \times q$ matrix:

$$
\int_{\mathcal{I}} \widehat{\Omega}_{T}^{1 / 2} \overline{\frac{\partial g^{\prime}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right)} \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \widehat{\Omega}_{T}^{1 / 2} d \omega_{2},
$$

and the matrix $D_{T}=\left[D_{1, T} \cdots D_{q, T}\right]$ and the diagonal matrix $\Lambda_{T}$ of eigenvalues $\lambda_{i, T}$ satisfy

$$
\int_{\mathcal{I}} \widehat{\Omega}_{T}^{1 / 2} \overline{\frac{\partial g^{\prime}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right)} \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \widehat{\Omega}_{T}^{1 / 2} d \omega_{2} D_{T}=D_{T} \Lambda_{T} .
$$

where $\widehat{\Omega}_{T}$ is a consistent estimate of $\Omega$.

We first show that $\left\|K_{T}-K\right\| \rightarrow 0$ in probability. Consider for a given element $(j, l)$ of $K_{T}\left(\omega_{1}, \omega_{2}\right)$ and the corresponding element of $K\left(\omega_{1}, \omega_{2}\right)$, the following expression:

$$
\int_{\mathcal{I}} \int_{\mathcal{I}} \widehat{k}_{j l, T}\left(\omega_{1}, \omega_{2}\right)-k_{j l}\left(\omega_{1}, \omega_{2}\right) d \omega_{1} d \omega_{2} .
$$

where $\widehat{k}_{j l, T}\left(\omega_{1}, \omega_{2}\right)=\frac{\partial g_{j}}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \widehat{\Omega} \overline{\frac{\partial g_{l}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right)^{\prime}}$ for $\widehat{\alpha}_{T}^{1}$ ia consistent first step estimator of $\alpha_{0}$ and $k_{j l}\left(\omega_{1}, \omega_{2}\right)=\frac{\partial g_{j}}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega_{1}\right) \Omega \overline{\frac{\partial g_{l}}{\partial \beta}\left(\alpha_{0}, \beta_{0}, \omega_{2}\right)^{\prime}}$. Under $\widehat{\alpha}_{T}^{1} \xrightarrow{p} \alpha_{0}, \widehat{\beta}_{T} \xrightarrow{p} \beta_{0}$ and $\widehat{\Omega}_{T} \xrightarrow{p} \Omega$, the expression above converges to zero. This holds for $\forall j, l=1, \ldots, J$ which implies that $\left\|K_{T}-K\right\| \rightarrow 0$ in probability. Let $\left(\widehat{K}_{T} f\right)\left(\omega_{1}\right)$ denote

$$
\left(\widehat{K}_{T} f\right)\left(\omega_{1}\right)=\left(\sum_{l=1}^{J} \int_{\mathcal{I}} \widehat{k}_{j l, T}\left(\omega_{1}, \omega_{2}\right) f_{l}\left(\omega_{2}\right) d \omega_{2}\right)_{j=1, \ldots, J}
$$

Then $\left(\widehat{K}_{T} f\right)\left(\omega_{1}\right)$ can be written as:

$$
\left(\widehat{K}_{T} f\right)\left(\omega_{1}\right)=\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \Omega_{T}^{1 / 2} \int_{\mathcal{I}} \Omega_{T}^{1 / 2} \overline{\frac{\partial g^{\prime}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right)} f\left(\omega_{2}\right) d \omega_{2}
$$

for $f(\omega)=\left(f_{1}(\omega), f_{2}(\omega), \ldots, f_{J}(\omega)\right)^{\prime}$ where $f_{j}(\omega)$ for $j=1, \ldots, J$ are scalar functions in $L^{2}(\mathcal{I})$. In this case, $R\left(K_{T}\right)$ is the space spanned by $\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \Omega_{T}^{1 / 2}$ with rank at most equals to $q$. The eigenfunctions $\gamma_{i, T}$ is necessarily of the form $\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \Omega_{T}^{1 / 2} D_{i, T}$ where the matrix $D_{i, T}$ is of dimension $q \times 1$ and $D_{T}=\left[\begin{array}{llll}D_{1, T} & D_{2, T} & \ldots & D_{q, T}\end{array}\right]$ where $D_{T}$ is of dimension $q \times q$. By virtue of the Mercer's theorem, the vector $\gamma_{i, T}(\omega)$ of eigenfunctions satisfies

$$
\left(K_{T} \gamma_{i, T}\right)\left(\omega_{1}\right)=\lambda_{i, T} \gamma_{i, T}\left(\omega_{1}\right)
$$

where $\lambda_{i, T}$ is the corresponding eigenvalue of the eigenfunctions vector $\gamma_{i, T}$.

Using $\gamma_{i, T}(\omega)=\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega\right) \Omega_{T}^{1 / 2} D_{i, T}$ yields:

$$
\left(K_{T} \gamma_{i, T}\right)\left(\omega_{1}\right)=\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \Omega_{T}^{1 / 2} \int_{\mathcal{I}} \Omega_{T}^{1 / 2} \overline{\frac{\partial g^{\prime}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right)} \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \Omega_{T}^{1 / 2} D_{i, T} d \omega_{2}
$$

Let $D_{T}=\left[\begin{array}{llll}D_{1, T} & D_{2, T} & \ldots & D_{q, T}\end{array}\right]$ and $\Lambda_{T}$ denote the matrices containing the eigenvectors and the eigenvalues of the following $q \times q$ matrix:

$$
\int_{\mathcal{I}} \Omega_{T}^{1 / 2} \overline{\frac{\partial g^{\prime}}{\partial \beta}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \Omega_{T}^{1 / 2} d \omega_{2} .
$$

More specifically, the eigenvectors $D_{i, T}, i=1, \ldots, q$ and the corresponding eigenvalues $\lambda_{i, T}$ solve the following system of $q$ equations:

$$
\int_{\mathcal{I}} \Omega_{T}^{1 / 2} \overline{\frac{\partial g^{\prime}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right)} \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \Omega_{T}^{1 / 2} d \omega_{2} D_{i, T}=\lambda_{i, T} D_{i, T} .
$$

Using the spectral decomposition,

$$
\begin{aligned}
&\left(K_{T} \gamma_{i, T}\right)\left(\omega_{1}\right)=\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \Omega_{T}^{1 / 2} \int_{\mathcal{I}} \Omega_{T}^{1 / 2} \frac{\overline{\partial g^{\prime}}}{\partial \beta}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \\
& \frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{2}\right) \Omega_{T}^{1 / 2} D_{T} d \omega_{2} \\
&=\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \Omega_{T}^{1 / 2} D_{i, T} \lambda_{i, T}=\lambda_{i, T} \gamma_{i, T}\left(\omega_{1}\right),
\end{aligned}
$$

which implies that $\gamma_{i, T}(\omega)$ is given by $\frac{\partial g}{\partial \beta^{\prime}}\left(\widehat{\alpha}_{T}^{1}, \widehat{\beta}_{T}, \omega_{1}\right) \Omega_{T}^{1 / 2} D_{i, T}$. A consistent estimator of the Moore-Penrose generalized inverse is then given by:

$$
\left(K_{T}^{-1} f\right)\left(\omega_{1}\right)=\sum_{i=1}^{q} \frac{1}{\lambda_{i, T}} \gamma_{i, T}\left(\omega_{1}\right)\left\langle f, \gamma_{i, T}\right\rangle
$$

### 1.3 Derivation of the first-step and second-step C-ALS estimator in the bivariate case

We show the following result.
Proposition 1.2. Consider the identifying restrictions that the effect of the second structural shock on the first variable in a bivariate structural VAR is zero over a frequency interval:

$$
\widehat{c}_{11}\left(e^{-i \omega}\right) a_{12}(0)+\widehat{c}_{12}\left(e^{-i \omega}\right) a_{22}(0)=0 \quad \forall \omega \in[\underline{\omega}, \bar{\omega}] .
$$

Then, using the simplified objective function, the optimal $C$-ALS is:

$$
\widehat{\tilde{a}}_{12, T}=-\frac{\widehat{s}_{11, T}^{\prime}\left(\widehat{W}_{2 T}^{2}\right)^{-1} \widehat{s}_{12, T}}{\widehat{s}_{11, T}^{\prime}\left(\widehat{W}_{2 T}^{2}\right)^{-1} \widehat{s}_{11, T}}
$$

where $\widehat{s}_{11, T}, \widehat{s}_{12, T}$ and $\widehat{W}_{2 T}$ are given by:

$$
\begin{aligned}
& \widehat{s}_{11, T}=\widehat{\Omega}^{\prime 1 / 2} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi} \widehat{\widetilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi}\right) \widehat{c}_{11, l}\right] \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega \\
& \widehat{s}_{12, T}=\widehat{\Omega}^{\prime 1 / 2} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi}\right) \widehat{c}_{12, l}\right] \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega
\end{aligned}
$$

where $\Phi=\operatorname{vec}\left(\Phi_{p}\right)$, the vectorization of the autoregressive parameters and

$$
\widehat{W}_{2 T}=\widehat{\Omega}^{\prime 1 / 2}\left[\sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi}\right)\left(\frac{\partial \widehat{c}_{11, l}}{\partial \Phi^{\prime}} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, l}}{\partial \Phi^{\prime}}\right) \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega\right] \widehat{\Omega}^{1 / 2}
$$

where $\widehat{\tilde{a}}_{12, T}^{1}$ is a first-step estimator and $\int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega=\frac{1}{l-j}[\sin ((l-j) \omega)]_{\underline{\omega}}^{\bar{\omega}}=\frac{2}{l-j} \sin ((l-j) \bar{\omega})$ with a symmetric interval $[-\bar{\omega}, \bar{\omega}]$ for $(l-j) \neq 0$ and $\int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega=\bar{\omega}-\underline{\omega}=2 \bar{\omega}$ for $l=j$.

Proof: The objective function of the C-ALS problem of a bivariate VAR model is based on the $W_{2}$ and $\underline{s}(\cdot, \cdot)$ matrices, with

$$
\begin{array}{r}
\underline{s}\left(\tilde{a}_{12}, \hat{\Phi}_{T}\right)=\int_{\underline{\omega}}^{\bar{\omega}} \sum_{j=0}^{\infty}\left[\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} e^{i \omega j} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} e^{i \omega j}\right) \Omega^{1 / 2}\right]^{\prime} \sum_{l=0}^{\infty}\left[\widehat{c}_{11, l} e^{-i \omega l} \tilde{a}_{12}+\widehat{c}_{12, l} e^{-i \omega l}\right] d \omega \\
=\int_{\underline{\omega}}^{\bar{\omega}} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\widehat{\tilde{a}}_{12, T}^{1} \frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{11, l} \tilde{a}_{12}+\left(\widehat{\tilde{a}}_{12, T}^{1} \frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{12, l}+\left(\frac{\partial \widehat{C}_{12, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{11, l} \tilde{a}_{12}+\left(\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{12, l}\right] \\
\times \exp ((j-l) \omega) d \omega \\
=\sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\widehat{\tilde{a}}_{12, T}^{1} \frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{11, l} \tilde{a}_{12}+\left(\widehat{\tilde{a}}_{12, T}^{1} \frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{12, l}+\left(\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{11, l} \tilde{a}_{12}+\left(\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} \Omega^{1 / 2}\right)^{\prime} \widehat{c}_{12, l}\right] \\
\times \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega
\end{array}
$$

and

$$
\begin{aligned}
\widehat{W}_{2 T} & =\int_{\underline{\omega}}^{\bar{\omega}} \sum_{j=0}^{\infty}\left[\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} e^{i \omega j} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} e^{i \omega j}\right) \Omega^{1 / 2}\right]^{\prime} \sum_{l=0}^{\infty}\left[\left(\frac{\partial \widehat{c}_{11, l}}{\partial \Phi^{\prime}} e^{-i \omega l} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, l}}{\partial \Phi^{\prime}} e^{-i \omega l}\right) \Omega^{1 / 2}\right] d \omega \\
& =\int_{\underline{\omega}}^{\omega} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}}\right) \Omega^{1 / 2}\right)^{\prime}\left(\left(\frac{\partial \widehat{c}_{11, l} \widehat{\tilde{a}}_{12, T}^{1}}{\partial \Phi^{\prime}}+\frac{\partial \widehat{c}_{12, l}}{\partial \Phi^{\prime}}\right) \Omega^{1 / 2}\right)\right] \exp ((j-l) \omega) d \omega \\
& =\sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}}\right) \Omega^{1 / 2}\right)^{\prime}\left(\left(\frac{\partial \widehat{c}_{11, l}}{\partial \Phi^{\prime}} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, l}}{\partial \Phi^{\prime}}\right) \Omega^{1 / 2}\right)\right] \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega
\end{aligned}
$$

and the last equality holds by the symmetry of the interval around zero. The objective function can be then rewritten as:

$$
\begin{equation*}
\underline{s}\left(\tilde{a}_{12}, \widehat{\Phi}_{T}\right)^{\prime}\left(\widehat{W}_{2 T}^{2}\right)^{-1} \underline{\underline{s}\left(\tilde{a}_{12}, \widehat{\Phi}_{T}\right)}=\left(\widehat{s}_{11, T} \tilde{a}_{12}+\widehat{s}_{12, T}\right)^{\prime}\left(\widehat{W}_{2 T}^{2}\right)^{-1}\left(\widehat{s}_{11, T} \tilde{a}_{12}+\widehat{s}_{12, T}\right) \tag{1.1}
\end{equation*}
$$

where $\widehat{s}_{11, T}, \widehat{s}_{12, T}$ and $\widehat{W}_{2 T}$ are given by:

$$
\begin{aligned}
& \widehat{s}_{11, T}=\sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\frac{\widehat{\tilde{a}}_{12, T}^{1} \partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \widehat{\Omega}^{1 / 2}\right)^{\prime} \widehat{c}_{11, l}+\left(\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} \widehat{\Omega}^{1 / 2}\right)^{\prime} \widehat{c}_{11, l}\right] \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega \\
& \widehat{s}_{12, T}=\sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\frac{\widehat{\tilde{a}}_{12, T}^{1} \partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \widehat{\Omega}^{1 / 2}\right)^{\prime} \widehat{c}_{12, l}+\left(\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}} \widehat{\Omega}^{1 / 2}\right)^{\prime} \widehat{c}_{12, l}\right] \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega,
\end{aligned}
$$

and

$$
\widehat{W}_{2 T}=\sum_{j=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\left(\frac{\partial \widehat{c}_{11, j}}{\partial \Phi^{\prime}} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, j}}{\partial \Phi^{\prime}}\right) \widehat{\Omega}^{1 / 2}\right)^{\prime}\left(\left(\frac{\partial \widehat{c}_{11, l}}{\partial \Phi^{\prime}} \widehat{\tilde{a}}_{12, T}^{1}+\frac{\partial \widehat{c}_{12, l}}{\partial \Phi^{\prime}}\right) \widehat{\Omega}^{1 / 2}\right)\right] \int_{\underline{\omega}}^{\bar{\omega}} \cos ((l-j) \omega) d \omega .
$$

The minimiser of the objective function 1.1 is given by $\widehat{\tilde{a}}_{12, T}=-\frac{\widehat{s}_{11, T}^{\prime}\left(\widehat{W}_{2 T}^{2}\right)^{-1} \widehat{s}_{12, T}}{\widehat{s}_{11, T}^{\prime}\left(\widehat{W}_{2 T}^{2}\right)^{-1} \hat{s}_{11, T}}$.

## 2 Misspecification and C-ALS

We start from the (functional) system of estimating equations (equation 4.17), which can be decomposed as follows:

$$
g\left(a(0), \beta_{0}, \omega\right)=\left[\begin{array}{c}
g_{1}\left(a(0), \beta_{0}\right)  \tag{2.2}\\
g_{2}\left(a(0), \beta_{0}, \omega\right)
\end{array}\right],
$$

where the estimating equations $g_{1}\left(a(0), \beta_{0}\right)=0$ result from vech $\left(\Sigma-A(0) A(0)^{\prime}\right)$ and the $q_{2}$ (functional) estimating equations $g_{2}\left(a(0), \beta_{0}, \omega\right)=0$, which stem from the frequency identifying restrictions, can be written as:

$$
\mathcal{G} \operatorname{vec}\left(A\left(e^{-i \omega}\right)-C\left(e^{-i \omega}\right) A(0)\right)=\mathbf{0}_{\mathbf{q}_{\mathbf{2}} \times \mathbf{1}}
$$

where $\mathcal{G}$ is an $q_{2} \times N^{2}$ selection matrix, $a(0)=\operatorname{vec}(A(0))$ and $\omega \in[\underline{\omega}, \bar{\omega}]$.
To analyze the impact of misspecification on our identification method, one can specify a local alternative of the estimating equations as follows:

$$
g_{2 T}\left(a(0), \beta_{0}, \omega\right)=g_{2}\left(a(0), \beta_{0}, \omega\right)+\frac{c}{\sqrt{T}} h(\omega)
$$

where $\frac{c}{\sqrt{T}} h(\omega)$ is the deviation from the null hypothesis, the scalar $c$ represents the distance from the null, and $h$ is a function of $\omega$, with $\omega \in[\underline{\omega}, \bar{\omega}]$, which gives the direction of the alternative hypothesis. Under the local alternative hypothesis, one can show that the asymptotic expansion of the C-ALS estimator is given by:

$$
\begin{aligned}
\sqrt{T}\left(\widehat{\alpha}_{T}-\alpha_{0}\right)= & -\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right)\right\rangle^{-1}\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), S \frac{\partial g}{\partial \beta^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right) \sqrt{T}\left(\widehat{\beta}_{T}-\beta_{0}\right)\right\rangle \\
& -\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right)\right\rangle^{-1}\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), c S H(\omega)\right\rangle \\
& +o_{p}(1)
\end{aligned}
$$

where the $H(\omega)$ matrix is partitioned comfortably w.r.t. (2.2), i.e.,

$$
H(\omega)=\left[\begin{array}{c}
0 \\
h(\omega)
\end{array}\right] .
$$

Due to the presence of the second right-hand side term, the C-ALS estimator is then asymptotically biased; the magnitude of the bias being a function of the local alternative.

In this respect, it can be shown that the asymptotic distribution of the $J_{T}$ statistic under the local alternative is given by:

$$
J_{T}=\left\|\sqrt{T} S_{T} g\left(\widehat{\alpha}, \widehat{\beta}_{T}, \omega\right)\right\|^{2} \rightarrow \sum_{j=1}^{m} \lambda_{j} \chi_{j}^{2}\left(1, \tilde{\delta}_{j}\right) .
$$

with the non-centrality parameter $\tilde{\delta}_{j}=c^{2}\left\langle S M H, \gamma_{j}\right\rangle$, and where the $\gamma_{j}$ terms are the eigenvectors of $S K S^{*}$ and $M=I-P$ is the operator from $H$ to $H$ such for all $f \in H$ :

$$
M f(\omega)=f(\omega)-\frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right)\left\langle S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right)\right\rangle^{-1}\left\langle S^{*} S \frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right), f(\omega)\right\rangle
$$

where $I$ denotes the identity operator on $H$, which is compact since $H$ is finite dimensional, and $P$ is the orthogonal projection onto the subspace of $\frac{\partial g}{\partial \alpha^{\prime}}\left(\alpha_{0}, \beta_{0}, \omega\right)$ with a range of dimension $r$. Accordingly, the $J_{T}$-statistic is asymptotically distributed as a sum of weighted noncentral $\chi^{2}$ distributions with 1 degree of freedom, and the non-centrality parameter is given by $\tilde{\delta}_{j}$, which is function of the local alternative.

As an illustration, using the data generating process of Section 5 and the same notations, we consider the following local alternative hypothesis:

$$
\begin{equation*}
\widehat{c}_{11}\left(e^{-i \omega}\right) a_{12}(0)+\widehat{c}_{12}\left(e^{-i \omega}\right) a_{22}(0)=\frac{c}{\sqrt{T}} h(\omega) \quad \forall \omega \in \mathcal{I}_{\omega}=[\underline{\omega}, \bar{\omega}] . \tag{2.3}
\end{equation*}
$$

For the first set of experiments in the paper, we assume that $\delta=.05$ as the alternative. Figure 1 reports the (partial) spectral density of the first variable with respect to the second structural shock as a measure of the deviation from the null hypothesis. It corresponds to the right-hand side term in eq. (2.3).

Figure 1: Deviation from the null hypothesis: second shock on first variable, $\delta=.05$


Figure 1 shows that the partial spectral density after imposing the local alternative is strongly concentrated around the zero frequency. This implies that considering a tight interval around zero would induce a significant bias of the estimator and a high probability of rejection for the $J$-test. When the frequency
band of interest, $\mathcal{I}_{\omega}$, gets larger and larger, our Monte Carlo simulations show that the bias decreases and the $J_{T}$-stat is less and less powerful, as to be expected.

As a second set of experiments, we consider a constant function $h$ on the interval $[-\pi, \pi]$. Then our simulation results show that the bias and the rejection rate of the $J_{T}$-stat are constant irrespective of the frequency band of interest $\mathcal{I}_{\omega}$ as expected. ${ }^{1}$

## 3 A discretized ALS estimator

One possible approach is to apply the standard asymptotic least squares procedure using a discretization of the frequency band, and thus evaluating $g\left(a_{0}, \widehat{\beta}_{T}, \omega_{\tau}\right)=0$ at different points/frequencies, say for $\tau=$ $1, \cdots, n$. We consider frequency intervals of the form $I_{\omega}=[-\bar{\omega} ;-\underline{\omega}] \cup[\underline{\omega} ; \bar{\omega}]$ where $\bar{\omega} \geq \underline{\omega} \geq 0$. In practise, only $[\underline{\omega} ; \bar{\omega}]$ needs to be considered for standard identifying restrictions $C(z) A(0)=0$.

### 3.1 Discretization of the frequency interval

We first divide the interval $[\underline{\omega} ; \bar{\omega}]$ into $n-1$ subintervals to obtain $n$ frequency indices. Let $\omega_{j}$ denote the $j$-th frequency in the partition:

$$
\underline{\omega}=\omega_{1}<\omega_{2}<\cdots<\omega_{n}=\bar{\omega} .
$$

On top of the estimating equations defined by vech $\left(\Sigma_{u}-A(0) A(0)^{\prime}\right)=0$, the $n$ frequency identifying restrictions write

$$
c_{11}\left(e^{-i \omega_{j}}\right) a_{12}(0)+c_{12}\left(e^{-i \omega_{j}}\right) a_{22}(0)=0
$$

### 3.2 First-step and second-step discretized ALS estimator

Proposition 3.1 provides the first-step and second-step discretized ALS estimator in the case of a bivariate $\operatorname{VAR}(\mathrm{p})$ model. ${ }^{2}$ Provided that the dimension of the system of estimating equations cannot exceed the dimension of the vector of auxiliary parameter $(2 n \leq \operatorname{dim}(\Phi))$, one can derive the second-step discretized ALS estimator.

Proposition 3.1. Consider a discretization of the frequency band

$$
\underline{\omega}=\omega_{1}<\omega_{2}<\cdots<\omega_{n}=\bar{w} .
$$

Suppose that $\left(X_{t}\right)$ is described by a bivariate $\operatorname{VAR}(p)$ model and that the identifying restriction is given by:

$$
\widehat{c}_{11}\left(e^{-i \omega}\right) \tilde{a}_{12}(0)+\widehat{c}_{12}\left(e^{-i \omega}\right)=0 .
$$

[^1]Then, the first-step discretized ALS estimator, denoted by $\widehat{\tilde{a}}_{12, T}^{1, d}$, is:

$$
\widehat{\tilde{a}}_{12, T}^{1, d}=-\frac{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty}\left(\widehat{c}_{11, k} \widehat{c}_{12, j} \sum_{j=1}^{n} \cos \left(\omega_{j}(k-\ell)\right)\right)}{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty}\left(\widehat{c}_{11, k} \widehat{c}_{11, j} \sum_{j=1}^{n} \cos \left(\omega_{j}(k-\ell)\right)\right)}
$$

and the second-step discretized ALS-estimator, denoted by $\widehat{\widetilde{a}}_{12, T}^{d}$, is:

$$
\widehat{\widetilde{a}}_{12, T}^{d}=-\frac{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \widehat{c}_{11, k} \widehat{c}_{12, \ell} \Lambda^{\prime}\left(\omega_{1: n}, k\right)\left[\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \alpha_{k, \ell}\left(\widehat{\widetilde{a}}_{12, T}^{1, d}\right) \Lambda^{\star}\left(\omega_{1: n}, k, \ell\right)\right]^{-1} \Lambda\left(\omega_{1: n}, \ell\right)}{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \widehat{c}_{11, k} \widehat{c}_{11, \ell} \Lambda^{\prime}\left(\omega_{1: n}, k\right)\left[\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \alpha_{k, \ell}\left(\widehat{\widetilde{a}}_{12, T}^{1, d}\right) \Lambda^{\star}\left(\omega_{1: n}, k, \ell\right)\right]^{-1} \Lambda\left(\omega_{1: n}, \ell\right)} .
$$

where $\alpha_{k, \ell}\left(\begin{array}{c}\widehat{a}_{12, T}^{1, d}\end{array}\right)=\left(\frac{\partial \widehat{c}_{11, k}}{\partial \beta^{\prime}} \widehat{\widetilde{a}}_{12, T}^{1, d}+\frac{\partial \widehat{c}_{12, k}}{\partial \beta^{\prime}}\right) \widehat{\Omega}_{T}\left(\frac{\partial \widehat{c}_{11, \ell}}{\partial \beta} \widehat{\widetilde{a}}_{12, T}^{1, d}+\frac{\partial \widehat{c}_{12, \ell}}{\partial \beta}\right)$ is a scalar, $\Lambda^{\star}\left(\omega_{1: n}, k, \ell\right):=\Lambda\left(\omega_{1: n}, k\right) \Lambda^{\prime}\left(\omega_{1: n}, \ell\right)$, and $\Lambda\left(\omega_{1: n}, k\right)=\left(\begin{array}{llllll}\cos \left(\omega_{1} k\right) & \cdots & \cos \left(\omega_{n} k\right) & \sin \left(\omega_{1} k\right) & \cdots & \sin \left(\omega_{n} k\right)\end{array}\right)^{\prime}$.

Proof: Following the approach of Feuerverger and McDunnough (1981), Singleton (2001) and Chacko and Viceira (2003), we distinguish the real part and the imaginary part of the identifying restrictions:

$$
\begin{aligned}
& \sum_{k=0}^{\infty} \widehat{c}_{11, k} \cos \left(\omega_{j} k\right) \tilde{a}_{12}(0)+\sum_{k=0}^{\infty} \widehat{c}_{12, k} \cos \left(\omega_{j} k\right)=0 \\
& \sum_{k=0}^{\infty} \widehat{c}_{11, k} \sin \left(\omega_{j} k\right) \tilde{a}_{12}(0)+\sum_{k=0}^{\infty} \widehat{c}_{12, k} \sin \left(\omega_{j} k\right)=0
\end{aligned}
$$

for $j=1, \cdots, n$. Accordingly, the moment conditions are given by:

$$
g\left(\tilde{a}_{12}(0), \widehat{\beta}_{T}, \omega_{1: n}\right)=\binom{g_{1}\left(\tilde{a}_{12}(0), \widehat{\beta}_{T}, \omega_{1: n}\right)}{g_{2}\left(\tilde{a}_{12}(0), \widehat{\beta}_{T}, \omega_{1: n}\right)}
$$

where

$$
g_{1}\left(\tilde{a}_{12}(0), \widehat{\beta}_{T}, \omega_{1: n}\right)=\left(\begin{array}{c}
\sum_{k=0}^{\infty} \widehat{c}_{11, k} \cos \left(\omega_{1} k\right) \tilde{a}_{12}(0)+\sum_{k=0}^{\infty} \widehat{c}_{12, k} \cos \left(\omega_{1} k\right) \\
\vdots \\
\sum_{k=0}^{\infty} \widehat{c}_{11, k} \cos \left(\omega_{n} k\right) \tilde{a}_{12}(0)+\sum_{k=0}^{\infty} \widehat{c}_{12, k} \cos \left(\omega_{n} k\right)
\end{array}\right)
$$

and

$$
g_{2}\left(\tilde{a}_{12}(0), \widehat{\beta}_{T}, \omega_{1: n}\right)=\left(\begin{array}{c}
\sum_{k=0}^{\infty} \widehat{c}_{11, k} \sin \left(\omega_{1} k\right) \tilde{a}_{12}(0)+\sum_{k=0}^{\infty} \widehat{c}_{12, k} \sin \left(\omega_{1} k\right) \\
\vdots \\
\sum_{k=0}^{\infty} \widehat{c}_{11, k} \sin \left(\omega_{n} k\right) \tilde{a}_{12}(0)+\sum_{k=0}^{\infty} \widehat{c}_{12, k} \sin \left(\omega_{n} k\right)
\end{array}\right)
$$

A first-step consistent estimator of $\tilde{a}_{12}(0)$ solves the following minimization problem (using the identity matrix of order $2 n$ ):

$$
\widehat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}=\underset{\tilde{a}_{12}}{\operatorname{argmin}} \quad g^{\prime}\left(\tilde{a}_{12}, \widehat{\beta}_{T}, \omega\right) g\left(\tilde{a}_{12}, \widehat{\beta}_{T}, \omega\right)
$$

or

$$
\widehat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}=\underset{\tilde{a}_{12}}{\operatorname{argmin}}\left\{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty}\left[\left(\widehat{c}_{11, k} \widehat{c}_{11, \ell} \tilde{a}_{12}^{2}+2 \widehat{c}_{11, k} \widehat{c}_{12, \ell} \tilde{a}_{12}+\widehat{c}_{12, k} \widehat{c}_{12, \ell}\right) \sum_{j=1}^{n} \cos \left(\omega_{j}(k-\ell)\right)\right]\right\} .
$$

Therefore,

$$
\widehat{\tilde{a}}_{12, \mathrm{~d}}^{1, \mathrm{~d}}=-\frac{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty}\left(\widehat{c}_{11, k} \widehat{c}_{12, \ell} \sum_{j=1}^{n} \cos \left(\omega_{j}(k-\ell)\right)\right)}{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty}\left(\widehat{c}_{11, k} \widehat{c}_{11, \ell} \sum_{j=1}^{n} \cos \left(\omega_{j}(k-\ell)\right)\right)} .
$$

Accordingly, the (discretized) second-step ALS estimator, denoted by $\widehat{\tilde{a}}_{12}^{\mathrm{d}}$, solves (provided that the dimension of the system of estimating equations does not exceed the dimension of the vector of auxiliary parameters):

$$
\widehat{\tilde{a}}_{12, T}^{\mathrm{d}}=\underset{\tilde{a}_{12}}{\operatorname{argmin}} g^{\prime}\left(\tilde{a}_{12}, \widehat{\beta}_{T}, \omega\right) S_{0}^{-1}\left(\hat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right) g\left(\tilde{a}_{12}, \widehat{\beta}_{T}, \omega\right)
$$

where $S_{0}^{-1}\left(\widehat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right)$ is the $2 n \times 2 n$ efficient weighting matrix defined by:

$$
S_{0}^{-1}\left(\widehat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right):=\left[\frac{\partial g\left(\hat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right)}{\partial \beta^{\prime}} \widehat{\Omega}_{T} \frac{\partial g^{\prime}\left(\hat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right)}{\partial \beta}\right]^{-1}
$$

with

$$
\begin{aligned}
& \frac{\partial g(.)}{\partial \beta^{\prime}} \widehat{\Omega}_{T} \frac{\partial g^{\prime}(.)}{\partial \beta}=\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \Lambda\left(\omega_{1: n}, k\right)\left(\frac{\partial \widehat{c}_{11, k}}{\partial \beta^{\prime}} \widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}+\frac{\partial \widehat{c}_{12, k}}{\partial \beta^{\prime}}\right) \widehat{\Omega}_{T}\left(\frac{\partial \widehat{c}_{11, \ell}}{\partial \beta} \widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}\right. \\
&=\sum_{k=0}^{\infty} \sum_{l=0}^{\infty}\left(\frac{\partial \widehat{c}_{12, \ell}}{\partial \beta}\right) \Lambda^{\prime}\left(\omega_{1: k}, \ell\right) \\
&\left.=\widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}+\frac{\partial \widehat{c}_{12, k}}{\partial \beta^{\prime}}\right) \widehat{\Omega}_{T}\left(\frac{\partial \widehat{c}_{11, \ell} \widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}}{\partial \beta}+\frac{\partial \widehat{c}_{12, \ell}}{\partial \beta}\right) \Lambda^{\star}\left(\omega_{1: n}, k, \ell\right) \\
& \alpha_{k, \ell}\left(\widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}\right) \Lambda^{\star}\left(\omega_{1: n}, k, \ell\right)
\end{aligned}
$$

where
$\Lambda^{\star}\left(\omega_{1: n}, k, \ell\right):=\left(\begin{array}{ccc|ccc}\cos \left(\omega_{1} k\right) \cos \left(\omega_{1} \ell\right) & \cdots & \cos \left(\omega_{1} k\right) \cos \left(\omega_{n} \ell\right) & \cos \left(\omega_{1} k\right) \sin \left(\omega_{1} \ell\right) & \cdots & \cos \left(\omega_{1} k\right) \sin \left(\omega_{n} \ell\right) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \cos \left(\omega_{n} k\right) \cos \left(\omega_{1} \ell\right) & \cdots & \cos \left(\omega_{n} k\right) \cos \left(\omega_{n} \ell\right) & \cos \left(\omega_{n} k\right) \sin \left(\omega_{1} \ell\right) & \cdots & \cos \left(\omega_{n} k\right) \sin \left(\omega_{n} \ell\right) \\ \hline \sin \left(\omega_{1} k\right) \cos \left(\omega_{1} \ell\right) & \cdots & \sin \left(\omega_{1} k\right) \cos \left(\omega_{n} \ell\right) & \sin \left(\omega_{1} k\right) \sin \left(\omega_{1} \ell\right) & \cdots & \sin \left(\omega_{1} k\right) \sin \left(\omega_{n} \ell\right) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sin \left(\omega_{n} k\right) \cos \left(\omega_{1} \ell\right) & \cdots & \sin \left(\omega_{n} k\right) \cos \left(\omega_{n} \ell\right) & \sin \left(\omega_{n} k\right) \sin \left(\omega_{1} \ell\right) & \cdots & \sin \left(\omega_{n} k\right) \sin \left(\omega_{n} \ell\right)\end{array}\right)$.

Note that the analytical expression of $\frac{\partial \widehat{c}_{11, j}{ }^{\prime}}{\partial \beta}$ and $\frac{\partial \widehat{c}_{12, j}}{\partial \beta^{\prime}}$ is provided in Appendix 4. In addition, it also worth noting that the rank of the matrix $\Lambda^{\star}\left(\omega_{1: n}, k, j\right)$ is one.

Under suitable regularity conditions, the discretized second-step ALS estimator, denoted $\widehat{\widetilde{a}}_{12, T}^{\mathrm{d}}$, solves the first-order condition:

$$
g^{\prime}\left(\widehat{\widetilde{a}}_{12}^{\mathrm{d}}, \widehat{\beta}_{T}, \omega_{1: n}\right) S_{0}^{-1}\left(\widehat{\tilde{a}}_{12, \mathrm{~d}}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right) \frac{\partial g\left(\widehat{\widetilde{a}}_{12, T}^{\mathrm{d}}, \widehat{\beta}_{T}, \omega_{1: n}\right)}{\partial \tilde{a}_{12}^{\prime}}=0
$$

that is,

$$
\left[\sum_{k=0}^{\infty}\left\{\left(\widehat{c}_{11, k} \widehat{\widetilde{a}}_{12, T}^{\mathrm{d}}+\widehat{c}_{12, k}\right) \Lambda^{\prime}\left(\omega_{1: n}, k\right)\right\}\right] S_{0}^{-1}\left(\widehat{\tilde{a}}_{12, T}^{1, \mathrm{~d}}, \widehat{\beta}_{T}, \omega_{1: n}\right)\left[\sum_{l=0}^{\infty} \widehat{c}_{11, \ell} \Lambda\left(\omega_{1: n}, \ell\right)\right]=0 .
$$

Finally,

$$
\widehat{\widetilde{a}}_{12, T}^{\mathrm{d}}=-\frac{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \widehat{c}_{11, k} \widehat{c}_{12, \ell} \Lambda^{\prime}\left(\omega_{1: n}, k\right)\left[\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \alpha_{k, \ell}\left(\widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}\right) \Lambda^{\star}\left(\omega_{1: n}, k, \ell\right)\right]^{-1} \Lambda\left(\omega_{1: n}, \ell\right)}{\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \widehat{c}_{11, k} \widehat{c}_{11, \ell} \Lambda^{\prime}\left(\omega_{1: n}, k\right)\left[\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \alpha_{k, \ell}\left(\widehat{\widetilde{a}}_{12, T}^{1, \mathrm{~d}}\right) \Lambda^{\star}\left(\omega_{1: n}, k, \ell\right)\right]^{-1} \Lambda\left(\omega_{1: n}, \ell\right)} .
$$

## 4 Identification of only one structural shock of interest

We also consider the identification of a single structural shock in a VAR with more than two variables without identifying other structural shocks. Without loss of generality, we assume that the structural shock of interest is the first one. For instance, this corresponds to the identification of a technology shock in a multivariate VAR without requiring the identification of other shocks (see Christiano et al., 2006b; Francis and Ramey, 2009) or the identification of a news shocks (Beaudry and Portier, 2006; Barsky and Sims, 2011; Kurmann and Sims, 2021). For sake of completeness, we first provide the common sign and (long-run) exclusion restrictions and then turn to the frequency identifying restrictions.

Following Christiano et al. (2006a), the dynamic effects of the first structural shock can be computed by identifying only the first column of $A(0),[A(0)] \cdot 1$, as combining a sign restriction and zero restrictions on the long-run impact uniquely identifies the vector $[A(0)] .{ }_{.1}$. In this respect, one needs to impose $N-1$ zero restrictions:

$$
\widehat{C}(1) A(0)=A(1)=\left[\begin{array}{cc}
a_{11}(1) & 0_{1 \times(N-1)} \\
\tilde{A}_{21}(1) & \tilde{A}_{22}(1)
\end{array}\right]
$$

where $\tilde{A}_{21}(1)$ is the first column of the long-run impact matrix $A(1)$ after dropping the first element $a_{11}(1)$, and the submatrix $\tilde{A}_{22}(1)$ contains the other columns of $A(1)$ (except the first row of those columns). Imposing that only the first structural shock has a long-run impact on the first variable yields the following
specification of the long-run variance-covariance matrix: ${ }^{3}$
$\widehat{C}(1) A(0) A(0)^{\prime} \widehat{C}(1)^{\prime}=A(1) A(1)^{\prime}=\left[\begin{array}{cc}a_{11}(1)^{2} & a_{11}(1) \tilde{A}_{21}(1)^{\prime} \\ \tilde{A}_{21}(1) a_{11}(1) & \tilde{A}_{21}(1) \tilde{A}_{21}(1)^{\prime}+\tilde{A}_{22}(1) \tilde{A}_{22}(1)^{\prime}\end{array}\right]=\widehat{C}(1) \widehat{\Sigma}_{T} \widehat{C}(1)^{\prime}$.
This implies that $a_{11}(1)^{2}$ is the $(1,1)$-element of the matrix $\widehat{C}(1) \widehat{\Sigma}_{T} \widehat{C}(1)^{\prime}$ and that $\tilde{A}_{21}(1)$ is equal to the corresponding elements of the matrix $\widehat{C}(1) \widehat{\Sigma}_{T} \widehat{C}(1)^{\prime}$ divided by $a_{11}(1)$. Because the first column of the matrix $C(1) A(0)$, denoted by $[A(1)]_{1}$, is known, the column vector $[A(0)]_{1}$ is uniquely identified by the relation $[A(0)]_{\cdot 1}=\widehat{C}(1)^{-1}[A(1)]_{\cdot 1}$. Now consider the same restrictions but in the frequency domain. One then has:

$$
\begin{aligned}
\left|a_{11}\left(e^{-i \omega}\right)\right|^{2} & =\left[\widehat{C}\left(e^{-i \omega}\right) \widehat{\Sigma}_{T} \overline{\widehat{C}\left(e^{-i \omega}\right)}\right]_{11} \\
\tilde{A}_{21}\left(e^{-i \omega}\right) a_{11}(z) & =\left[\widehat{C}\left(e^{-i \omega}\right) \widehat{\Sigma}_{T}{\overline{\widehat{C}}\left(e^{-i \omega}\right)}^{\prime}\right]_{n 1, n=2, \ldots, N}
\end{aligned}
$$

where $\left[\widehat{C}\left(e^{-i \omega}\right) \widehat{\Sigma}_{T}{\overline{\widehat{C}}\left(e^{-i \omega}\right)}^{\prime}\right]_{n 1}$ is the element $(n, 1)$ of the matrix $\widehat{C}\left(e^{-i \omega}\right) \widehat{\Sigma}_{T}{\widehat{\widehat{C}}\left(e^{-i \omega}\right)}^{\prime}$. Then it is straightforward to show the following result.
Proposition 4.1. Consider the following identifying restrictions $\sum_{j=1}^{N} \widehat{c}_{1 j}\left(e^{-i \omega}\right) a_{j n}(0)=0$ for $n=2, \ldots, N$ and $\forall \omega \in[\underline{\omega}, \bar{\omega}]$. Let $\widehat{\beta}_{T}=\left(\operatorname{vec}\left(\widehat{\Phi}_{p}\right)^{\prime} \text {, vech }\left(\widehat{\Sigma}_{T}\right)^{\prime}\right)^{\prime}$ denote the vector of dimension $q=N^{2} \times p+\frac{N(N+1)}{2}$ of the reduced-form parameters estimates. Under Assumptions A. 1 and A.2, the estimating equations, $g\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right)=\mathbf{0}$ for $\omega \in[\underline{\omega}, \bar{\omega}]$, defined by

$$
g\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right)=\left(g_{1}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right), g_{2}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right), \cdots, g_{N}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right)\right)^{\prime}
$$

with

$$
\begin{aligned}
& g_{1}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right)=\left|\left[\widehat{C}\left(e^{-i \omega}\right) A(0)\right]_{11}\right|^{2}-\left[\widehat{C}\left(e^{-i \omega}\right) \widehat{\Sigma}_{T}{\overline{\widehat{C}}\left(e^{-i \omega}\right)}^{\prime}\right]_{11} \\
& g_{n}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right)=\left[\widehat{C}\left(e^{-i \omega}\right) A(0)\right]_{n 1}\left[A(0)^{\prime}{\overline{\widehat{C}}\left(e^{-i \omega}\right)}^{\prime}\right]_{11}-\left[\widehat{C}\left(e^{-i \omega}\right) \widehat{\Sigma}_{T}{\overline{\widehat{C}\left(e^{-i \omega}\right)}}^{\prime}\right]_{n 1}
\end{aligned}
$$

for $n=2, \ldots, N$ uniquely identify the first column of the matrix $\alpha_{0}=[A(0)] \cdot 1$ up to a sign restriction.
Note that for $n=1, \cdots, N$ the moment conditions can be written as:

$$
g_{n}\left(\alpha_{0}, \widehat{\beta}_{T}, \omega\right)=\sum_{r=1}^{N} \sum_{s=1}^{N} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} c_{1 s, j} c_{n r, l}\left(a_{r 1}(0) a_{s 1}(0)-\widehat{\sigma}_{s r, T}\right) \cos ((j-l) \omega)
$$

where $\widehat{\sigma}_{s r, T}$ is a consistent estimate of the $(s, r)$ element of $\Sigma$. As in Section 5.1, a first-step consistent C-ALS estimator of $\alpha_{0}=[A(0)] .1$ solves the following minimization problem (using the identity operator as a weighting matrix):

$$
\widehat{\alpha}_{T}=\underset{a}{\arg \min } \int_{\underline{\omega}}^{\bar{\omega}} g\left(\alpha, \widehat{\beta}_{T}, \omega\right)^{\prime} \overline{g\left(\alpha, \widehat{\beta}_{T}, \omega\right)} d \omega
$$

[^2]Then the second-step C-ALS results from the minimization of the simplified (regularized) objective function in light of Proposition 3.2.. In particular, the first-order conditions can be derived from the following proposition.

Proposition 4.2. The first-order partial derivatives of the estimating equations,

$$
g\left(a, \hat{\beta}_{T}, \omega\right)=\operatorname{vech}\left(\widehat{C}(z) \widehat{\Sigma}_{T} \widehat{C}^{*}(z)-\widehat{C}(z) A(0) A(0)^{\prime} \widehat{C}^{*}(z)\right)
$$

with respect to $a$, $\Phi$, and $\sigma$ are respectively given by:

$$
\frac{\partial g}{\partial a^{\prime}}\left(a, \hat{\beta}_{T}, \omega\right)=-L_{N}(\bar{C}(z) \otimes C(z))\left(I_{N^{2}}+K_{N N}\right)\left(A(0) \otimes I_{N}\right)
$$

and
$\frac{\partial g}{\partial \Phi^{\prime}}\left(a, \hat{\beta}_{T}, \omega\right)=L_{N}\left[\left(I_{N} \otimes \widehat{C}(z)\left(\widehat{\Sigma}_{T}-A(0) A(0)^{\prime}\right)\right) \frac{\partial v e c\left(\widehat{C}^{*}(z)\right)}{\partial \Phi^{\prime}}+\left(\widehat{\widehat{C}}(z)\left(\widehat{\Sigma}_{T}-A(0) A(0)^{\prime}\right) \otimes I_{N}\right) \frac{\partial v e c(\widehat{C}(z))}{\partial \Phi^{\prime}}\right]$
$\frac{\partial g}{\partial \sigma^{\prime}}\left(a, \hat{\beta}_{T}, \omega\right)=L_{N}(\bar{C}(z) \otimes C(z))$
with $\widehat{C}^{* \prime}(z)=\widehat{\widehat{C}(z)}, L_{N}$ is an $\left(\frac{1}{2} N(N+1) \times N^{2}\right)$ elimination matrix, $K_{N N}$ is the commutator matrix for which $K_{N N} \operatorname{vec}(X)=\operatorname{vec}\left(X^{\prime}\right)$ and $X$ is an arbitrary $N \times N$ matrix.

Proof: One has

$$
\begin{aligned}
\frac{\partial g}{\partial a^{\prime}}\left(a, \hat{\beta}_{T}, \omega\right) & =-L_{N} \frac{\partial}{\partial a^{\prime}} \operatorname{vec}\left(\widehat{C}(z) A(0) A(0)^{\prime} \widehat{C}^{*}(z)\right) \\
& =-L_{N}(\widehat{\widehat{C}}(z) \otimes \widehat{C}(z)) \frac{\partial}{\partial a^{\prime}} \operatorname{vec}\left(A(0) A(0)^{\prime}\right) \\
& =-L_{N}(\overline{\widehat{C}}(z) \otimes \widehat{C}(z))\left[\left(I_{N} \otimes A(0)\right) \frac{\partial \operatorname{vec}\left(A(0)^{\prime}\right)}{\partial a^{\prime}}+\left(A(0) \otimes I_{N}\right) \frac{\partial \operatorname{vec}(A(0))}{\partial a^{\prime}}\right] \\
& =-L_{N}(\bar{C}(z) \otimes C(z))\left(\left(I_{N} \otimes A(0)\right) K_{N N}+\left(A(0) \otimes I_{N}\right)\right) \\
& =-L_{N}(\bar{C}(z) \otimes C(z))\left(I_{N^{2}}+K_{N N}\right)\left(A(0) \otimes I_{N}\right)
\end{aligned}
$$

using that $K_{N N}\left(A(0) \otimes I_{N}\right)=\left(I_{N} \otimes A(0)\right) K_{N N}$. On the other hand, the partial derivatives with respect to $\Phi$ are obtained using the standard product rule for vector differentiation with the vec operator:

$$
\frac{\partial \mathrm{vec}(A(\theta) C D(\theta))}{\partial \theta^{\prime}}=\left(I_{N} \otimes A(\theta) C\right) \frac{\partial \mathrm{vec}(D(\theta))}{\partial \theta^{\prime}}+\left(D(\theta)^{\prime} C^{\prime} \otimes I_{N}\right) \frac{\partial \mathrm{vec}(A(\theta))}{\partial \theta^{\prime}}
$$

Moreover, $\frac{\partial \operatorname{vec}(C(z))}{\partial \Phi^{\prime}}$ can be derived from Lütkepohl (2007, p. 111). Finally, using the property $\operatorname{vec}\left(P Q P^{*}\right)=$ $\left(P^{\star \prime} \otimes P\right) \operatorname{vec}(Q)$ and $P^{\star \prime}=\bar{P}$ where $P$ is a complex-valued matrix, one has

$$
\frac{\partial g}{\partial \sigma^{\prime}}\left(a, \hat{\beta}_{T}, \omega\right)=L_{N}(\bar{C}(z) \otimes C(z)) \frac{\partial \mathrm{vec}(\Sigma)}{\partial \sigma^{\prime}}=L_{N}(\bar{C}(z) \otimes C(z))
$$

## 5 Other Monte Carlo simulations

In this section, we provide some further Monte Carlo simulations to study the finite sample performances of the C-ALS estimator. We still assume that the data generating process (DGP) is a bivariate VAR model in
which the first variable, $X_{1, t}$, is nonstationary and thus written in first-difference and the second variable, $X_{2, t}$, is a weakly stationary process:

$$
\begin{align*}
\Delta X_{1, t} & =\vartheta_{1}+\rho_{11,1} \Delta X_{1, t-1}+\left(\rho_{12,1}+\delta\right) X_{2, t-1}-\rho_{12,1} X_{2, t-2}+\epsilon_{1, t}  \tag{5.4}\\
X_{2, t} & =\vartheta_{2}+\rho_{21,1} \Delta X_{1, t-1}+\rho_{22,1} X_{2, t-1}+\rho_{22,2} X_{2, t-2}+b_{21} \epsilon_{1, t}+\epsilon_{2, t}, \tag{5.5}
\end{align*}
$$

where the vector $\epsilon_{t}=\left(\epsilon_{1, t}, \epsilon_{2, t}\right)^{\prime}$ represents some structural shocks, with $\epsilon_{t} \sim N\left(0, I_{2}\right)$.

As second set of experiments, we consider other parameter configurations. For instance, as shown in the online Appendix, using a $\operatorname{VAR}(2)$ specification with $\left(\rho_{11,1}, \rho_{12,1}, \rho_{21,1}, \rho_{22,1}, \rho_{22,2}, b_{21}\right)=(0,-0.08,0.2, \rho+$ $0.55,-0.55 \rho, 0.2)$ under the null $(\delta=0)$ and alternative $(\delta>0)$ hypothesis where $\rho=0.9,0.95$, or 0.98 , our results are qualitatively similar.

Looking at Figures 2 to 4 , the results are qualitatively similar to those of the VAR(1) under the null hypothesis. Interestingly, when $\delta=0$, the contribution of the second structural shock to the first variable is not zero but is rather close to zero and decreases when the length of the frequency band increases, and, as such, this case can be interpreted as a local alternative to the null hypothesis of a well-specified identifying restriction. ${ }^{4}$ Given that this local alternative hypothesis is often considered as a very plausible DGP in SVARs applications, the finite sample properties of the second-step C-ALS estimator are again remarkable and appealing. Meanwhile, as to be expected, the J-stat has less power but can still provide useful information in the case of a local alternative.

When $\delta=0.04$ and there is a misspecified exclusion restriction (alternative hypothesis), three points are worth emphasizing in Figures 5 to 7 . First, the C-ALS estimator still dominates the first-step, MS, and LR approaches and displays a small (cumulative) mean absolute bias for intervals greater than $\omega_{30}$, whereas the max-share estimator is only slightly improving relative to the standard LR approach. Notably, the mean absolute bias of the two-step C-ALS estimator is close to zero for the widest interval $\omega_{30}$, while it increases with decreasing intervals. This arises as the variance contribution of the second shock to the first variable augments when the frequency interval becomes smaller and smaller, and the (regularized) minimisation problem of the C-ALS estimator seeks to find the optimal linear combination of the reduced-form shocks such that the contribution of the second structural shock to the first variable is minimized. Second, the bias reduction of the second-step C-ALS estimator is achieved with a lower (cumulative) RMSE relative to other estimators. Third, all these results are robust irrespective of the chosen horizon $H$.

[^3]
## 6 An overview of the spectral density matrix and its application to SVAR models

The second-order properties of a second-order stationary, bivariate, zero-mean time series $\left\{X_{t}\right\}_{-\infty<t<+\infty}$ may be described either by the auto- and cross-covariances $\Gamma(h)=E\left(X_{t+h} X_{t}^{\prime}\right)$ or, equivalently, by the spectral density matrix or spectrum $f(\omega)$ of the process, which is given by:

$$
f_{X}(\omega)=\frac{1}{2 \pi}\left[\begin{array}{ll}
f_{X, 11}(\omega) & f_{X, 12}(\omega) \\
f_{X, 21}(\omega) & f_{X, 22}(\omega)
\end{array}\right], \quad \omega \in[-\pi ; \pi]
$$

Provided that $\sum_{h=-\infty}^{+\infty}\left|\gamma_{i j}(h)\right|<\infty$ where $\gamma_{i j}(h)$ is the element $(\mathrm{i}, \mathrm{j})$ of the matrix $\Gamma(h)$, the marginal spectral densities $f_{X, i i}(\omega)$ and the cross-spectrum $f_{X, i j}(\omega)$ are defined as the Fourier transforms of the auto- and cross-covariance functions: ${ }^{5}$

$$
f_{X, i j}(\omega)=\frac{1}{2 \pi} \sum_{h=-\infty}^{+\infty} \gamma_{i j}(h) e^{-i \omega h}
$$

By the inverse of the Fourier transform, one retrieves the auto- and cross-covariances as

$$
\gamma_{i j}(h)=\int_{-\pi}^{\pi} e^{i \omega h} f_{X, i j}(\omega) d \omega
$$

In particular, when $h=0$, the spectrum integrates to the unconditional variance-covariance matrix of $X_{t}$,

$$
\Gamma(0)=\int_{-\pi}^{\pi} f_{X}(\omega) d \omega
$$

Said differently, marginal spectral densities and the cross-spectrum integrate to the unconditional variances and the unconditional covariance between $X_{1}$ and $X_{2}$. The marginal spectral density (respectively, the cross-spectrum) at frequency $\omega$ is the portion of the variance (respectively, covariance between $X_{1}$ and $X_{2}$ ) that is attributable to cycles with frequency $\omega$. More generally, for any $\omega$ between 0 and $\pi$, the expression

$$
\int_{-w}^{w} f_{X}(\omega) d \omega
$$

provides the decomposition of the unconditional variance-covariance matrix that is attributable to the frequency interval $[0 ; \omega]$. In this respect, the spectral matrix is a natural tool to analyze fluctuations of macro variables at different periodicities and, in particular, at business cycle, medium-term, and low (or long-run) frequencies. ${ }^{6}$

[^4]In the case of the moving average representation in terms of innovations of the reduced-form VAR(p) (equation 2.1 of the main text), the spectral density matrix is given by:

$$
\begin{equation*}
f_{X}(\omega)=\frac{1}{2 \pi} C\left(e^{-i \omega}\right) \Sigma_{u} C^{\star}\left(e^{-i \omega}\right) \quad \omega \in[-\pi ; \pi], \tag{6.6}
\end{equation*}
$$

where $C\left(e^{-i \omega}\right)=\sum_{k=0}^{\infty} C_{k} e^{-i \omega k}$. In the case of the moving average representation in terms of structural shocks (equation 2.3 of the main text), one has:

$$
\begin{align*}
f_{X}(\omega) & =\frac{1}{2 \pi} A\left(e^{-i \omega}\right) A^{\star}\left(e^{-i \omega}\right) \quad \omega \in[-\pi ; \pi]  \tag{6.7}\\
& =\frac{1}{2 \pi} C\left(e^{-i \omega}\right) A(0) A(0)^{\prime} C^{\star}\left(e^{-i \omega}\right) .
\end{align*}
$$

Using equations (6.6) and (6.7), we unsurprisingly recover the mapping (equation 2.5 of the main text) between the reduced-form moving average matrix $C\left(e^{-i \omega}\right)$ and the structural moving average matrices $A(0)$ and $A\left(e^{-i \omega}\right)$ for a given frequency $\omega$ :

$$
C\left(e^{-i \omega}\right) A(0)=A\left(e^{-i \omega}\right)
$$

Thus, imposing identifying restrictions in the frequency domain can be achieved by imposing constraints on some elements of $C\left(e^{-i \omega}\right) A(0)$. These constraints yield estimating equations at a given frequency $\omega$, and, more generally, on a frequency interval.

## 7 Comparison between C-ALS and the CEV-based approach

We first describe the approach of Christiano, Eichenbaum and Vigfusson (2006a, 2006b), and, without loss of generality, we consider the case of a just-identified bivariate structural VAR model. Then we reformulate their nonparametric correction using a system of estimating equations. Finally, we discuss the main differences.

### 7.1 Approach of Christiano, Eichenbaum and Vigfusson (2006a, 2006b)

(a) Derivation of $a(0)$ In the case of a just-identified bivariate structural VAR model, the approach of Christiano, Eichenbaum and Vigfusson (2006a, 2006b) rests on the following equations:

$$
\begin{aligned}
& A(0)=C(1)^{-1} A(1) \equiv \Phi(1) A(1) \\
& \frac{1}{2 \pi} A(1) A(1)^{\prime}=\frac{1}{2 \pi} C(1) \Sigma_{u} C(1)^{\prime}=f_{X}(0)
\end{aligned}
$$

where $f_{X}(0)$ is the spectral matrix at frequency $\omega=0$. At this stage, we assume that $f_{X}(0)$ is given and we discuss later on the use of a nonparametric correction of the estimate of $f_{X}(0)$. Using the second equation and assuming that $A(1)$ is a lower-triangular matrix (that is, one imposes the identifying restriction $\left.[A(1)]_{12}=0\right)$, it is straightforward to show that:

$$
\left(\begin{array}{cc}
a_{11}^{2}(1) & a_{11}(1) a_{21}(1) \\
a_{11}(1) a_{21}(1) & a_{21}^{2}(1)+a_{22}^{2}(1)
\end{array}\right)=\left(\begin{array}{cc}
S_{11}(0) & S_{21}^{\prime}(0) \\
S_{21}(0) & S_{22}(0)
\end{array}\right)
$$

where $S_{X}(0)=2 \pi f_{X}(0)$. Using the sign restrictions $a_{11}(1) \geq 0$ and $\left.a_{22}(1) \geq 0\right)$, one can easily get the expressions of $a_{11}(1), a_{21}(1)$, and $a_{22}(1)$ as some functions of the elements of $S_{X}(0)$ :

$$
\begin{aligned}
& a_{11}^{2}(1)=S_{11}(0) \Leftrightarrow a_{11}(1)=\sqrt{S_{11}(0)} \\
& a_{21}(1)=\frac{S_{21}^{\prime}(0)}{a_{11}(0)}=\frac{S_{21}^{\prime}(0)}{\sqrt{S_{11}(0)}} \\
& a_{22}^{2}(1)=S_{22}(0)-\frac{S_{21}^{\prime 2}(0)}{S_{11}(0)}(\geq 0) \Leftrightarrow a_{22}(1)=\sqrt{S_{22}(0)-\frac{S_{21}^{\prime 2}(0)}{S_{11}(0)}} .
\end{aligned}
$$

Now, using the first equation $A(0)=C^{-1}(1) A(1)=\Phi(1) A(1)$, one has

$$
\left(\begin{array}{ll}
a_{11}(0) & a_{12}(0) \\
a_{21}(0) & a_{22}(0)
\end{array}\right)=\left(\begin{array}{cc}
\Phi_{11}(1) & \Phi_{12}(1) \\
\Phi_{21}(1) & \Phi_{22}(1)
\end{array}\right)\left(\begin{array}{cc}
\sqrt{S_{11}(0)} & 0 \\
\frac{S_{21}^{\prime}(0)}{\sqrt{S_{11}(0)}} & \sqrt{S_{22}(0)-\frac{S_{21}^{\prime 2}(0)}{S_{11}(0)}}
\end{array}\right) .
$$

where the second right-hand side term is the lower triangular Cholesky factor of $S_{X}(0)$. Therefore,

$$
\begin{aligned}
& a_{11}(0)=\Phi_{11}(1) \sqrt{S_{11}(0)}+\Phi_{12}(1) \frac{S_{21}^{\prime}(0)}{\sqrt{S_{11}(0)}} \\
& a_{12}(0)=\Phi_{12}(1) \sqrt{S_{22}(0)-\frac{S_{21}^{\prime 2}(0)}{S_{11}(0)}} \\
& a_{21}(0)=\Phi_{21}(1) \sqrt{S_{11}(0)}+\Phi_{22}(1) \frac{S_{21}^{\prime}(0)}{\sqrt{S_{11}(0)}} \\
& a_{22}(0)=\Phi_{22}(1) \sqrt{S_{22}(0)-\frac{S_{21}^{\prime 2}(0)}{S_{11}(0)}}
\end{aligned}
$$

It is worth emphasizing that the expression of $\tilde{a}_{12}(0)$, which is given by

$$
\tilde{a}_{12}(0)=\frac{a_{12}(0)}{a_{22}(0)}=\frac{\Phi_{12}(1)}{\Phi_{22}(1)} \equiv-\frac{C_{12}(1)}{C_{11}(1)},
$$

does not depend on $S_{X}(0)$.
(b) Nonparametric correction Consider now a local average estimator of $f_{X}(0)$ using the standard Andrews-Monahan estimator: ${ }^{7}$

$$
\widehat{f}_{X}(0)=\frac{1}{2 \pi} \widehat{C}(1) \widehat{F}_{u}(0) \widehat{C}(1)^{\prime}
$$

where $\widehat{F}_{u}(0)=\frac{1}{T} \sum_{k=-r}^{r}\left|1-\frac{k}{r}\right| \sum_{t=k}^{T} \widehat{u}_{t} \widehat{u}_{t-k}^{\prime}$.

[^5]Accordingly, the previous results are still valid after replacing $S_{X}(0)$ by $\widehat{S}_{X}(0)=2 \pi \widehat{f}_{X}(0)$, and $\Phi_{i j}(1)$ by some consistent estimate $\widehat{\Phi}_{i j}(1)$ for $i, j=1,2$. Notably, it turns out that the estimate of $\tilde{a}_{12}(0)$ is the same as the standard LR estimate, i.e. the nonparametric correction is only effective for the elements of $A(0)$, but not for $\tilde{a}_{12}(0)$.

### 7.2 Reformulation

Noting that $C(1) A(0)=S_{X, \operatorname{tr}}(0)$ where $S_{X, \operatorname{tr}}$ is the lower-triangular matrix that results from the Cholesky decomposition of the spectral matrix at $\omega=0$ (since there is no imaginary part) conveys the same information as the standard long run identifying restriction $c_{11}(1) a_{12}(0)+c_{12}(1) a_{22}(0)=0$, it turns out that the system of estimating equations of Christiano, Eichenbaum and Vigfusson (2006a, 2006b) can be written as:

$$
\begin{aligned}
& c_{11}(1) a_{12}(0)+c_{12}(1) a_{22}(0)=0 \\
& \operatorname{vech}\left(A(0) A(0)^{\prime}\right)=\operatorname{vech}\left(\Phi(1) S_{X}(0) \Phi(1)^{\prime}\right)
\end{aligned}
$$

whereas the standard long-run (LR) estimator is defined from:

$$
\begin{aligned}
& c_{11}(1) a_{12}(0)+c_{12}(1) a_{22}(0)=0 \\
& \operatorname{vech}\left(A(0) A(0)^{\prime}\right)=\operatorname{vech}\left(\Sigma_{u}\right)
\end{aligned}
$$

In contrast, the C-ALS estimator rests on:

$$
\begin{aligned}
& c_{11}\left(e^{-i \omega}\right) a_{12}(0)+c_{12}\left(e^{-i \omega}\right) a_{22}(0)=0 \\
& \operatorname{vech}\left(A(0) A(0)^{\prime}\right)=\operatorname{vech}\left(\Sigma_{u}\right)
\end{aligned}
$$

for all $\omega \in[\underline{\omega} ; \bar{\omega}]$.

On the one hand, using the Andrews-Monahan (1992) estimate of $f_{X}(0)$, one has

$$
\operatorname{vech}\left(A(0) A(0)^{\prime}-\widehat{\Phi}(1) \widehat{C}(1) \widehat{F}_{u}(0) \widehat{C}(1)^{\prime} \widehat{\Phi}(1)^{\prime}\right)=\mathbf{0}
$$

where $\widehat{\Phi}(1) \widehat{C}(1)=\mathrm{I}_{2}$ (since $\left.\widehat{C}(1)=\widehat{\Phi}^{-1}(1)\right)$ and $\widehat{F}_{u}(0)$ is defined above. Therefore, after plugging in a consistent estimate of the slope parameters and the residuals of the reduced-form VAR, the system of estimating equations is:

$$
\begin{aligned}
& \widehat{c}_{11}(1) a_{12}(0)+\widehat{c}_{12}(1) a_{22}(0)=0 \\
& \operatorname{vech}\left(A(0) A(0)^{\prime}-\widehat{F}_{u}(0)\right)=0
\end{aligned}
$$

where

$$
\widehat{F}_{u}(0) \simeq \widehat{\Sigma}_{T}+\sum_{\substack{k=-r \\ k \neq 0}}^{r}\left|1-\frac{k}{r}\right| \widehat{\Sigma}_{T}(k)
$$

where $\widehat{\Sigma}_{T}(k)$ is a consistent estimate of the k-order autovariance matrix of innovations.

### 7.3 Discussion

The previous system of estimating equations can be first compared with the one derived from the LR estimator:

$$
\begin{aligned}
& \widehat{c}_{11}(1) a_{12}(0)+\widehat{c}_{12}(1) a_{22}(0)=0 \\
& \operatorname{vech}\left(A(0) A(0)^{\prime}-\widehat{\Sigma}_{T}\right)=0
\end{aligned}
$$

Comparing the two systems of estimating equations, the correction of the spectral matrix at $\omega=0$ does not convey so much "new" information for the elements of $A(0)$ when $\widehat{\Sigma}_{T} \simeq \widehat{F}_{u}(0)$, i.e. when the reduced-form VAR is not misspecified (lag order) and that there is nothing left in the residuals (innovations). In presence of a misspecified VAR, the correction might be more substantial with respect to the standard LR estimator of $A(0)$.

On the other hand, the key point is that the approach of Christiano, Eichenbaum and Vigfussion (2006a, 2006b) makes use of a "local average" of the spectrum at $\omega=0$, but left unchanged the identifying constraint, which has a zero Lebesgue measure. Indeed, they consider a generic estimator of the true spectral matrix $f_{0}$ with the following form:

$$
\widehat{f}_{X}(0)=\int_{-\pi}^{\pi} W_{M}(\omega) I_{n}(\omega) d \omega
$$

where $W_{M}(\omega)$ is the periodogram window corresponding to the weights of the lag window $w(x)$, M is the truncation lag such that $M \rightarrow \infty$ and $M / n \rightarrow 0$, and $I_{n}(\omega)$ is the periodogram of the mean corrected data or residuals:

$$
\begin{aligned}
I_{n}(\omega) & =(2 \pi n)^{-1}\left[\Gamma_{z}(0)+2 \sum_{j=1}^{n-1} \Gamma_{z}(j) e^{-i \omega j}\right] \\
W_{M}(\omega) & =(2 \pi)^{-1}\left[w(0)+2 \sum_{j=1}^{n-1} w(j / M) e^{-i \omega j}\right]
\end{aligned}
$$

where $z_{t}=X_{t}-\mu_{X}$ or $z_{t}=\widehat{u}_{t}$ for all $t$, and $\Gamma_{z}(j)$ is the autocovariance matrix of order $j$. Then it can be shown (Hauser et al., 1999; Pötscher, 2002) that alternatives $f_{1}$ close to $f_{0}$ can have values $f_{1}(0)$ that are arbitrarily far away from the true value of the spectral density matrix at $\omega=0$. It is a direct consequence of the highly discontinuous nature of the functional $f_{X} \rightarrow f_{X}(0)$ with respect to the $L_{1}$-distance (see Section 8 of this Appendix). In this respect, the nonparametric correction proposed by Christiano, Eichenbaum and Vigfussion (2006a, 2006b) is not immune to the unreliability problem and still falls in the category of "ill-posed" (identification) problems (Sims, 1972).

## 8 Applications of the C-ALS approach

This section provides some applications of the asymptotic least squares theory using a frequency band or equivalently a continuum of estimating equations. The first two applications, namely the identification
of technology and news shocks, are further discussed respectively in the main text and in Section ?? of this technical appendix. The third application regarding neutral versus investment technology shocks is presented formally in Section 5. The fourth application is related to the recoverability condition of Chahrour and Jurado (2021) and the structural identification of expectations-driven fluctuations. The last application, that is, the identification and inference of common features, is left for future work. Finally, it is worth emphasizing that this section is not exhaustive in the sense that our methodology applies for general zero restrictions, i.e. with a mixture of short-run, delayed and long-run restrictions (e.g., in the monetary policy model of Rubio-Ramirez et al. (2010)), as well as for structural VARs based on present value models (e.g., Campbell and Shiller $(1987,1988)$ ), for structual VECM (e.g., King et al. (1991)) for the identification and contribution of seasonal cycles versus business cyles (Wen, 2002), and the estimation and validation of dynamic stochastic general equilibrium models using SVARs, among others.

### 8.1 The hours-productivity debate using bivariate SVAR models

The predominant role of technology shock as the main source behind movements in macro data has been sharply challenged since the important contribution of Galí (1999). Indeed, bivariate structural vector autoregressive models including labor productivity and hours worked yield conflicting results regarding the effect of technology shocks on hours worked, generally due to the assumed data generating process for the measure of hours worked (in level or in difference). On the other hand, Francis and Ramey (2009) show that demographic trends and sectoral allocation are important sources of low-frequency movements in hours worked and labor productivity. Consequently, labor productivity might be driven by two permanent shocks, the technology shock and the demographic shock, and thus the usual long-run restriction of hours-productivity VAR models might be violated. A SVAR model with some identifying restrictions on a frequency band allows to focus in a neighborhood more or less close to $\omega=0$ rather than just the zero frequency, and thus to assess the effects of other low-frequency movements, such as those advocated by Francis and Ramey (2005), on the identification of technology shocks.

Our strategy has three advantages with respect to the usual long-run restrictions. First, the set of identifying restrictions in the frequency domain has a Lebesgue measure strictly greater than zero. Indeed the fact that standard long-run restrictions (when $\omega=0$ ) have a zero Lebesgue measure (Faust, 1996) leads to the so-called unreliability problem (Faust and Leeper, 1997). Notably, unreliable long-run effects of shocks are transferred on estimates of other model parameters through the long-run identification scheme and thus any test of the null hypothesis that the $k$ th coefficient of an autoregressive polynomial in the SVAR equals zero is not consistent, i.e., the test has significance level greater than or equal to maximum power. In addition, one cannot compute asymptotically correct confidence intervals for impulse responses since the unreliability of the long run effect estimator is transferred to the estimator of the dynamic multipliers of the structural shocks. In contrast, and as suggested by Faust and Lepper (1997), Faust (1998) and Pötscher (2002), this issue can be circumvented by imposing restrictions on the (long-run) effect of these shocks at non-zero frequencies and not only at the zero-frequency so that the problem is no-longer ill-posed in the terminology
of Pötscher (2002). Second, the overidentification testing procedure proposed in the next section allows to assess the hypothesis that only the first shock drives the long-run movements in labor productivity (Francis and Ramey, 2009). Third, using a wider frequency band relative to the zero frequency should help to better identify and estimate the structural shocks driving long-run movements.

### 8.2 News shock

Using structural VAR's and partial identification schemes, recent empirical literature delivered controversial results concerning the role of anticipated neutral technology-news-shocks in business cycle fluctuations. By imposing long-run restrictions, Beaudry and Portier (2006) and Beaudry and Lucke (2010) conclude that news shocks about future productivity are one of the main drivers of business cycles and there is a positive (contemporaneous) impact of the news shock on hours worked. These results have been challenged in several dimensions. For instance, adopting partial identification schemes based on different max-share approaches, Barsky and Sims (2011) and Kurmann and Sims (2021) find results incompatible with the news-driven explanation of business cycles and thereby more in line with the implications of the standard neoclassical framework. These alternative identification strategies are based on the forecast error variance decomposition over a horizon of up to 40 quarters (Barsky and Sims, 2011) or 80 quarters (Kurmann and Sims, 2021): both strategies encompassing short-run and business cycle fluctuations. In contrast, using an appropriate frequency band has the advantage to focus on the medium and long-run frequencies of TFP and thus allows to isolate the identification of the news shocks from the effects of short-run and business cycles fluctuations. In addition we can test whether TFP is driven by one structural shock (the news shock) or two structural shocks (the surprise TFP shock and the news shock) in the medium to long-run. In the latter, any linear combination of the two structural shocks would be a main driver and thus there is a lack of proper identification.

### 8.3 Neutral versus investment-related technology shocks

Fisher (2006) examines the relative importance of neutral technology shocks and the investment-related technology shocks in the explanation of business cycles by incorporating long-run restrictions that separately identify these two sources of technology shocks. ${ }^{8}$ Pursuing this decomposition, Chen and Wemy (2015) argue that long-run movements in the capital-producing sector can spread and spillover to the rest of the economy and enhance TFP in long-run. Accordingly, the long-run fluctuations of TFP may be characterized by two stochastic trends driven unequivocally either by the long-run movements of specific TFP or the spillover effect of long-run movements due to investment-specific technological (IST) changes. This implies that long-run movements of the TFP series would be caused by two shocks while long-run movements of investment-specific technology are only driven by its own (structural) shock. In this respect, consider a SVAR in which the first variable is the IST series and the second variable is the TFP series of Fernald (2014). The two restrictions

[^6]that only one shock has long-run effects on IST, and the same shock and the neutral technology shocks have long-run effects on TFP lead to the following restrictions for an interval of frequencies around zero:
\[

C\left(e^{-i \omega}\right) A(0)=A\left(e^{-i \omega}\right)=\left[$$
\begin{array}{ccc}
a_{11}\left(e^{-i \omega}\right) & 0 & 0_{1 \times(N-2)} \\
a_{21}\left(e^{-i \omega}\right) & a_{22}\left(e^{-i \omega}\right) & 0_{1 \times(N-2)} \\
\tilde{A}_{31}\left(e^{-i \omega}\right) & \tilde{A}_{32}\left(e^{-i \omega}\right) & \tilde{A}_{33}\left(e^{-i \omega}\right)
\end{array}
$$\right]
\]

where $\tilde{A}_{31}\left(e^{-i \omega}\right)$ is the first column of the matrix $A\left(e^{-i \omega}\right)$ after dropping the two first elements $a_{11}\left(e^{-i \omega}\right)$ and $a_{21}\left(e^{-i \omega}\right)$ and the column vector $\tilde{A}_{32}\left(e^{-i \omega}\right)$ is the corresponding second column and the submatrix $A_{33}\left(e^{-i \omega}\right)$ contains the other columns of $A\left(e^{-i \omega}\right)$ (except the two first rows of those columns). Proceeding with a frequency band in a neighboorhood of $\omega=0$ allows to perform a statistical test on the assumption that one stochastic trend against the alternative hypothesis of two stochastic trends drives the long-run movements of TFP.

### 8.4 Recoverability and expectations-driven fluctuations

Chahrour and Jurado (2021) propose an identification condition of structural shocks, which is less restrictive than the usual condition of fundamentalness-the so-called recoverability. This condition only imposes that the structural shocks can be recovered from the past, present and future observables available to the econometrician. Indeed, the econometrician has access to the entire sample to identify the structural shocks and not only to the information available in the observables up to time $t$ as required by fundamentalness. Said differently, the econometrician can also use available observables at time $t+1, \ldots, t+h$ to infer structural shocks at time $t$. In this respect, the necessary and sufficient condition of recoverability depends on the invertibility of the Fourier transform of the two-side moving average representation of the observable variables as function of the structural shocks.

Consider the example in Section 3 of Chahrour and Jurado (2021) regarding the identification of a noise and a fundamental structural shock about technology as potential drivers of business cycle fluctuations. Their VAR procedure can be summarized as follows. On the one hand, the VAR estimation is conducted with a set of observables/variables including a technology measure $\left(a_{t}\right)$. On the other hand, one can get the joint spectral density of technology $a_{t}$ as well as the optimal forecast of technology $b_{t}$ implied by the VAR. Finally, the identification of the structural shocks is then achieved by matching the joint spectral density resulting from the VAR and the joint spectral density implied by the structural two-side moving average of $a_{t}$ and $b_{t}$ as function of the two structural shocks for the interval $\omega \in[-\pi, \pi]$. In our framework, this can be written as:

$$
g(\alpha, \widehat{\beta}, \omega)=\operatorname{vech}\left(\widehat{C}_{a b}\left(e^{-i \omega}\right) \widehat{\Sigma} \widehat{C}_{a b}\left(e^{-i \omega}\right)^{*}-A_{a b}\left(e^{-i \omega}\right) A_{a b}\left(e^{-i \omega}\right)^{*}\right)
$$

for $\omega \in[-\pi, \pi]$ where $\widehat{C}_{a b}\left(e^{-i \omega}\right) \widehat{\Sigma} \widehat{C}_{a b}\left(e^{-i \omega}\right)^{*}$ is the estimated joint spectral density of the technology series $a_{t}$ and its optimal forecast $b_{t}$ implied by the VAR, and $A_{a b}\left(e^{-i \omega}\right) A_{a b}\left(e^{-i \omega}\right)^{*}$ is the structural joint spectral density compatible with the identification restrictions. Then Chahrour and Jurado (2021) solves a minimization problem by using the factorization proposed by Rozanov (1967), and thus can derive the mapping from
observables to structural disturbances, as well as the impulse response functions and the (historical) variance decomposition. An optimal two-step C-ALS estimator can then be performed allowing to formally test the imposed identification restrictions that lead to recover the noise and the fundamental technology shocks. In the event of rejection, the measure of technology $a_{t}$ could be contaminated by measurement errors or shortrun fluctuations such that the identification restrictions hold only for medium-run and long-run frequencies. The C-ALS framework is well-suited to investigate such a conjecture.

### 8.5 Common features in the frequency domain

Since the seminal contribution of Engle and Kozicki (1993), a common feature can be defined as follows: a feature is common if a group of variables of interest possesses this feature and a combination of these variables does not have the feature. Canonical examples include cointegration in which some (all) variables have stochastic trends but some linear combinations of these variables do not have stochastic trends; common serial correlation in which some linear combinations of serially correlated variables correspond to a weak white noise process (Engle and Kozicki, 1993); common cycles in which linear combinations of the cycle components of a group of variables have no cyclical component (Vahid and Engle, 1993, and Hecq, Palm and Urbain, 2006). ${ }^{9}$ It turns out that this concept of a common feature can be extended to the existence of common business cycles, i.e. a set of series is characterized by some common business cycle fluctuations whereas some linear combinations does not have this feature. ${ }^{10}$

To go one step further, suppose that the number of structural shocks is less than the number of variables and thus that there exist some common business cycles, say on a frequency band $\omega \in[\underline{\omega}, \bar{\omega}]$. In this case, it turns out that the matrix $C\left(e^{-i \omega}\right)$ have less than full rank for all $\omega \in[\underline{\omega}, \bar{\omega}]$, meaning that the left null space of the matrix $C\left(e^{-i \omega}\right)$ is non-empty:

$$
C\left(e^{-i \omega}\right) A(0)=A\left(e^{-i \omega}\right)=\left[\begin{array}{ll}
A_{1}\left(e^{-i \omega}\right) & A_{2}\left(e^{-i \omega}\right)_{N \times s}
\end{array}\right]=\left[\begin{array}{ll}
A_{1}\left(e^{-i \omega}\right) & 0_{N \times s}
\end{array}\right] .
$$

Accordingly, there exists a set of $s$ linear independent combinations such that the rank of the matrix $A\left(e^{-i \omega}\right)$ is equal to $N-s$ and thus this rank restriction allows the identification of a subset of structural shocks. Moreover, since these linear combinations define a set of estimating equations, an overidentification test (see Proposition 3.3) can be conducted and interpreted as a reduced-rank test of common business cycles.

## 9 Unreliability of the long-run identification scheme

We discuss the unreliability of the long-run identification scheme. It can be explained from these two fundamental relationships:

$$
\begin{equation*}
C(1) A(0)=A(1), \tag{9.8}
\end{equation*}
$$

[^7]and
\[

$$
\begin{equation*}
A(L)=C(L) A(0)=C(L) C(1)^{-1} A(1), \tag{9.9}
\end{equation*}
$$

\]

It turns out that the long-run identification scheme conducts to reliable inference if and only if the $A(1)$ is consistently estimated in finite samples and especially the lag order $p$ is not misspecified. Otherwise, any inconsistent estimate of $A(1)$ leads to unreliable long-run effects of shocks (in finite samples). This in turn is transferred to the estimates of the dynamic multipliers of the structural shocks by virtue of Eq. (9.9). ${ }^{11}$ In particular, one cannot form asymptotically correct confidence intervals for impulse responses of each structural shock and there is no consistent test that an individual impulse response coefficient is zero (Faust and Leeper, 1997). The fundamental issue is that the true data generating process may have an infiniteordered VAR representation with $\Phi_{0}(L)=\sum_{j=1}^{\infty} \Phi_{0, j} L^{j}$ and thus the infinite sequence $\Phi_{0}=\left\{\Phi_{0,1}, \Phi_{0,2}, \cdots\right\}$ must be approximated by a finite sequence $\widetilde{\Phi}_{p}=\left\{\Phi_{1}, \cdots, \Phi_{p}\right\}$ (i.e., a misspecified VAR model). Such finite-parameter approximations to infinite lag distributions have been studied extensively by Sims (1971, 1972) and Pötscher (2002), especially for least-squares criterion. ${ }^{12}$ An accurate approximation from the point of view of least-squares fit does not imply an accurate approximation of the long run effect. ${ }^{13}$ This means that convergence of the sequence $\tilde{\Phi}_{p}$ is not sufficient to guarantee the convergence of some functions of those parameters (Sims, 1971,1972; Pötscher, 2002) as pointwise convergence does not imply (locally) uniform convergence. More specifically, functions of a lag distribution (e.g., the sum of coefficients) are in general discontinuous with respect to the metric implied be least-squares estimation. ${ }^{14}$ Say differently, the best least-squares approximation of $\Phi_{0}, \Phi_{p}$, might be arbitrarily close (w.r.t. $L_{2}$-norm) whereas $\Phi(1)$ and $\tilde{\Phi}_{p}(1)$ are arbitrarily far apart and thus converge to different limits. This stems also from the fact that the least-squares criterion at a single frequency admits a zero Lebesgue measure. From a practical point of view, it turns out that standard errors of estimates or the coefficient of determination might approach their optimum values in arbitrarily large samples while the estimated sum of coefficients remains arbitrarily far from their true values. Inference based on the sum of coefficients is then highly unreliable unless $\Phi$ is in fact contained in $\Phi_{p}$, and not only close to it (Pötscher, 2002). ${ }^{15}$

[^8]
## References

[1] Angeletos, G-M., Collard, F. and H. Dellas (2020), "Business-Cycle Anatomy", American Economic Review, vol. 110(10), 3030-3070.
[2] Barsky, R. B., and E. R. Sims (2011), "News Shocks and Business Cycles", Journal of Monetary Economics, vol. 58(3), 273-289.
[3] Beaudry, P., and B. Lucke (2010), "Letting Different Views about Business Cycles Compete", in NBER Macroeconomics Annual 2009, Volume 24, NBER Chapters, pp. 413-455. National Bureau of Economic Research, Inc.
[4] Beaudry, P., and F. Portier (2004), "An Exploration into Pigou's Theory of Cycles", Journal of Monetary Economics, vol. 51(6), 11831216.
[5] Beaudry, P., and F. Portier (2006), "Stock Prices, News, and Economic Fluctuations", American Economic Review, vol. 96(4), 1293-1307.
[6] Beaudry, P., and F. Portier (2014), "News-driven business cycles: Insights and challenges", Journal of Economic Literature, vol. 52(4), 993-1074.
[7] Ben Zeev, N. and H. Khan (2015), "Investment-specific News Shocks and U.S. Business Cycles", Journal of Money, Credit, and Banking, vol. 47(7), 1443-1464.
[8] Chacko, G. and L. Viceira (2003), "Spectral GMM estimation of continuous-time processes", Journal of Econometrics, vol. 116, 259-292.
[9] Campbell, J.Y., and R.J. Shiller (1987), "Cointegration and Tests of Present Value Models", Journal of Political Economy, vol. 95, 1062-1088.
[10] Campbell, J.Y., and R.J. Shiller (1988), "Stock Prices, Earnings and Expected Dividends", Journal of Finance, vol. 43, 661-676.
[11] Chahrour, R. and K. Jurado (2021), "Recoverability and Expectations-Driven Fluctuations", Review of Economic Studies, forthcoming.
[12] Chari, V., Kehoe, P., and E. McGrattan (2008), "Are Structural VARs with long-run Restrictions Useful in Developing Business Cycle Theory?", Journal of Monetary Economics, vol. 55(8), 1337-1352.
[13] Chen, K. and E. Wemy (2015), "Investment-specific Technological Changes: The source of Long-run TFP Fluctuations", European Economic Review, vol. 80, 230-252.
[14] Christiano, L.J., and R.J. Vigfusson (2003), "Maximum Likelihood in the Frequency Domain: The Importance of Time-to-Plan", Journal of Monetary Economics, vol. 50(4), 789-815.
[15] Christiano, L.J., Eichenbaum, M. and C.L. Evans (2005),"Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", Journal of Political Economy, vol. 113 (1), 1-45
[16] Christiano, L.J., Eichenbaum, M. and R. Vigfusson (2006a), "Assessing Structural VARs", NBER Macroeconomics Annual, vol. 21., 1-106.
[17] Christiano, L.J., Eichenbaum, M. and R. Vigfusson (2006b), "Alternative Procedures for Estimating Vector Autoregressions Identified with Long-run Restrictions", Journal of the European Economic Association, vol. 4(2-3), 475-483.
[18] Diebold, F.X., Ohanian, L.E. and J. Berkowitz (1998),," Dynamic Equilibrium Economies: A Framework for Comparing Models and Data", Review of Economic Studies, vol. 65(3), 433-451.
[19] Engle, R.F. and S. Hylleberg (1996), "Common Seasonal Features: Global Unemployment", Oxford Bulletin of Economics and Statistics, vol. 58, 615-630.
[20] Engle, R.F. and S. Kozicki (1993), "Testing for Common Features", Journal of Business \& Economic Statistics, vol. 11(4), 369-380.
[21] Erceg, Guerrieri and Gust (2005), "Can Long-run Restrictions Identify Technology Shocks?", Journal of the European Economic Association, vol. 3(6), 1237-1278.
[22] Faust, J. (1996), "Near Observational Equivalence and Theoretical Size Problems with Unit Root Tests", Econometric Theory, vol. 12, 724-731.
[23] Faust, J. (1998), "The Robustness of Identified VAR Conclusions about Money", Carnegie-Rochester Conference Series ob Public Policy, vol. 49(0), 207-44.
[24] Faust, J., and E.M. Leeper (1997), "When Do Long-Run Identifying Restrictions Give Reliable Results?", Journal of Business \& Economic Statistics, vol. 15, 345-353.
[25] Feuerverger, A. and P. McDunnough (1981), "On Some Fourier Methods for Inference", J. R. Statist. Asso., vol. 76, 379-397.
[26] Fisher, J.D.M. (2006), "The Dynamic Effects of Neutral and Investment-specific Technology Shocks", Journal of Political Economy, vol. 114(3), 413-451.
[27] Francis, N., and V. Ramey (2009), "Measures of Per Capita Hours and Their Implications for the Technology-Hours Debate", Journal of Money, Credit and Banking, vol. 41(6), 1071-1097.
[28] Francis, N., and V. Ramey (2005), "Is the Technology-driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited", Journal of Monetary Economics, vol. 52, 1379-1399.
[29] Galí, J. (1999), "Technology, Employment and the Business Cycle: So Technology Shocks Explain Aggregate Productivity?", American Economic Review, vol. 41, 1201-1249.
[30] Gouriéroux, C. and I. Peaucelle (1992), "Séries codépendantes", Actualité Economique, vol. 68(1-2), 283-304.
[31] Hendry, D.F. (1999), "A Theory of Co-Breaking", in Forecasting Non-Stationary Economic Time Series, eds. Hendry D.F. and P. Clements, Camridge, MIT Press.
[32] Hecq, A., Palm, F.C. and J-P. Urbain (2006), "Common Cyclical Features Analysis in VAR Models with Cointegration", Journal of Econometrics, vol. 132, 117-141.
[33] King, R., Plosser, C., Stock, J. and M. Watson (1991), "Stochastic Trends and Economic Fluctuations", American Economic Review, vol. 81(4), 819-40.
[34] Kurmann, A. and E. Sims (2021), "Revisions in Utilization-Adjusted TFP and Robust Identification of News Shocks", The Review of Economics and Statistics, vol. 103(2), 216235
[35] Newey, W. K. and D. McFadden, (1994), "Large sample estimation and hypothesis testing", Handbook of Econometrics, in: R. Engle \& D. McFadden (ed.) volume 4, chapter 36, Elsevier.
[36] Phillips, P.C.B. (1998), "Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs", Journal of Econometrics, vol. 83, 21-56.
[37] Pötscher, B.M. (2002), "Lower Risk Bounds and Properties of Confidence Sets for Ill-Posed Estimation Problems with Applications to Spectral Density and Persistence Estimation, Unit Roots, and Estimation of Long Memory Parameters", Econometrica, 70, 1035-1065.
[38] Qu, Z. and D. Tkachenko (2012a), "Identification and Frequency Domain Quasimaximum Likelihood Estimation of Linearized Dynamic Stochastic General Equilibrium Models", Quantitative Economics, vol. 3(1), 95-132
[39] Qu, Z. and D. Tkachenko (2012b), "Frequency Domain Analysis of Medium Scale DSGE Models with Application to Smets and Wouters (2007)", Advances in Econometrics, vol. 28: DSGE Models in Macroeconomics-Estimation, Evaluation and New Developments, Emerald, 319-385.
[40] Ramey, V.A. (2016), Macroeconomic Shocks and Their Propagation, edited in Handbook of Macroeconomics.
[41] Rotemberg, J.J., and M. Woodford (1999), "The Cyclical Behavior of Prices and Costs", NBER Working Papers 6909, National Bureau of Economic Research.
[42] Rozanov, Y.A. (1967), Stationary random processes, Holden-Day, San Francisco.
[43] Sala, L. (2015), "DSGE Models in the Frequency Domain", Journal of Applied Econometrics, vol. 30, 219-240.
[44] Sims, C. (1971), "Distributed Lag Estimation When the parameter Space is Explicitly InfiniteDimensional", Annals of Mathematical Statistics, vol. 42, 1622-1636.
[45] Sims, C. (1972), "The Role of Approximation Prior Restrictions in Distributed Lag Estimation", Journal of the American Statistical Association, vol. 67, 169-175.
[46] Singleton, K. J. (2001), "Estimation of Affine Pricing Models Using the Empirical Characteristic Function", Journal of Econometrics, vol. 102, 111-141.
[47] Vahid, F. and R.F. Engle (1993), "Common Trends and Common Cycles", Journal of Applied Econometrics, vol. 8, 341-360.

Figure 2: Contemporaneous bias and RMSE using a $\operatorname{VAR}(2)$ model with $\rho=.95$ and $\delta=0$


[^9]Figure 3: Cumulative Bias and RMSE up to 12 quarters using a $\operatorname{VAR}(2)$ model with $\rho=.95$ and $\delta=0$


Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent the LR, second-step C-ALS, Max-share and first-step estimators, respectively.

Figure 4: Impulse Responses for the first shock on second variable with $n=60, \rho=.95$ and $\delta=0$


Note: Confidence intervals are based on the 95-percentile from 10,000 Monte-Carlo experiments.

Figure 5: Contemporaneous bias and RMSE at the impact using a $\operatorname{VAR}(2)$ model with $\rho=.95$ and $\delta=.04$


[^10]Figure 6: Cumulative Bias and RMSE up to 12 quarters using a $\operatorname{VAR}(2)$ model with $\rho=.95$ and $\delta=.04$




Second shock on second variable: bias



Frequencies


First shock on second variable: RMSE


Second shock on second variable: RMSE


[^11]Figure 7: Impulse Responses for the first shock on second variable with $n=60, \rho=.95$ and $\delta=.04$


Note: Confidence intervals are based on the 95-percentile from 10,000 Monte-Carlo experiments.


[^0]:    *Université du Québec à Montréal, Chaire en macroéconomie et prévisions ESG-UQAM, and CIREQ, e-mail: guay.alain@uqam.ca.
    ${ }^{\dagger}$ EDHEC Business School, e-mail: florian.pelgrin@edhec.edu.

[^1]:    ${ }^{1}$ Simulations results are available upon request.
    ${ }^{2}$ Results can be easily generalizable in the case of any N -variate VAR(p) model.

[^2]:    ${ }^{3}$ Note that many matrices $A(0)$ are conformable with these restrictions, but the first column of each of these matrices $[A(0)] \cdot 1$ is the same (see Christiano et al., 2006b).

[^3]:    ${ }^{4}$ In contrast to other cases, the partial spectral density of the first variable with respect to the second structural shock is U-shaped in a neighborhood of $\omega=0$.

[^4]:    ${ }^{5}$ The real part of the cross-spectrum is the co-spectrum, whereas the imaginary part is the quadrature spectrum.
    ${ }^{6}$ Stationary time series are generally split into three ranges in the frequency domain. Indeed, the general consensus is that high-frequency or business-cycle frequencies correspond to periodicity below 8 years, whereas medium-term frequencies are those with periodicity in between 8 and 40-50 years and low or long-run frequencies are those with periodicity above 40-50 years.

[^5]:    ${ }^{7}$ One could also consider the Bartlett's nonparametric estimator of the population spectrum. Results are available upon request.

[^6]:    ${ }^{8}$ See Ramey (2016) for a survey on the debate on the relative importance of the neutral TFP shocks and the investment specific technology shocks. See also Ben Zeev and Khan (2015) for the effect of IST news shocks.

[^7]:    ${ }^{9}$ Other common features have been proposed in the literature which include common seasonally Engle and Hylleberg, 1996), codependence (Gouriéroux and Peaucelle, 1992), common structural breaks (Hendry, 1999) among others.
    ${ }^{10}$ Angeletos, Collard and Dellas (2020) find some support for a main business-cycle driver which implies that the business cycles fluctuations can be explained by a small number of structural shocks.

[^8]:    ${ }^{11}$ Using Monte Carlo simulations, Erceg et al. (2005) and Chari et al. (2008) study the extent of these small-sample estimation problems.
    ${ }^{12} \mathrm{~A}$ similar argument can be found in Christiano et al. (2006a).
    ${ }^{13}$ See Faust $(1996,1999)$ for an application of this result to unit root tests and confidence intervals for points on spectrum.
    ${ }^{14}$ The functional $S_{\Delta X} \rightarrow S_{\Delta X}(0)$, with $S_{\Delta X}$ the spectrum of the stochastic process $\left(\Delta X_{t}\right)$, is highly discontinuous w.r.t.
    $L_{2}$-distance. This makes the problem fall into the category of ill-posed problem (Sims, 1972; Pötscher, 2002).
    ${ }^{15}$ Note that it might occur regardless of how large the sample size is.

[^9]:    Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent the LR, second-step C-ALS, Max-share and first-step estimators, respectively.

[^10]:    Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent the LR, second-step C-ALS, Max-share and first-step estimators, respectively.

[^11]:    Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent the LR, second-step C-ALS, Max-share and first-step estimators, respectively.

