Estimating a VAR with R *

Alessia Paccagnini[a] Fabio Parla[b]

[a]University College Dublin & CAMA

[b]Central Bank of Ireland

SER - International Seminar on Statistics with R

June 2021

*The views expressed in these notes are those of the authors and do not necessarily reflect the views of the Central Bank of Ireland or the ESCB. Any errors are our own.

Macroeconomics and Reality, Econometrica (1980)

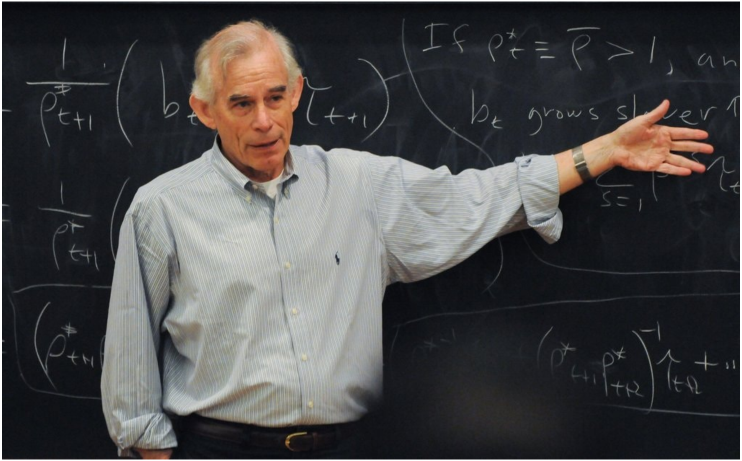

Figure 1: Christopher A. Sims, Nobel Memorial Prize in Economic Sciences in 2011

Outline

1. Vector Autoregressions (VAR)
2. VAR estimation
3. Lag length selection
4. Structural VAR
5. Application: illustrative example and R codes

Vector Autoregressive (VAR) model

- Econometric models largely used in empirical research.
- VARs model time series data.
- They are used to address questions like:
  - How much of an increase of a shock will affect the business cycle in the next years?
  - What is the contribution of a shock on the observed fluctuations in business cycle over time?

Some references

- Sims, Christopher A., 1980. Macroeconomics and reality. Econometrica, 1-48.
- Sims, Christopher A., Stock, James H., and Watson, Mark W., 1990. Inference in linear time series models with some unit roots. Econometrica, 113-144.
- Stock, James H., and Watson, Mark W., 2001. Vector autoregressions. Journal of Economic perspectives 15(4), 101-115.
- Kilian, Lutz, and Lütkepohl, Helmut, 2017. Structural vector autoregressive analysis. Cambridge University Press.
- Blake, Andrew P., and Mumtaz, Haroon, 2017. Applied Bayesian econometrics for central bankers. Bank of England, Centre for Central Banking Studies

VAR specification

- Multivariate extension of univariate autoregressive (AR) models.
- K-dimensional vector of endogenous variables, y_t, is described (over time) as a function of its past realizations $(y_{t-1}, \dots)$ and a vector of stochastic terms (u_t).
- A VAR with a lag of order p, named VAR(p):

$$y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + c + u_t \quad , \quad \text{with} \quad u_t \sim \mathcal{N}(0, \Sigma)$$

where:

$y_t = K \times 1$ vector of endogenous variables

$A_\ell = K \times K$ coefficients matrix, for $\ell = 1, \ldots, p$

$c = K \times 1$ vector of constant terms

- Other deterministic terms, such as time trends, and/or exogenous variables can be included. Here we skip for the sake of simplicity.

VAR residuals and variance-covariance matrix

- The VAR residuals, u_t, is a $K \times 1$ vector of zero-mean white noise process with a non singular covariance matrix, Σ, which is not assumed to be diagonal:

$$u_t \sim \mathcal{N}(0, \Sigma)$$

$$E(u_t u_s') = \begin{cases} \Sigma & \text{if } t = s \\ 0 & \text{if } t \neq s \end{cases}$$

- Σ is not diagonal: the residuals of the K equations might be contemporaneously correlated.

Example

Road map

- Consider $K = 2$ endogenous variable: $y_{1,t}$ and $y_{2,t}$ entering a VAR(1):
- 1^{st} specification: Vector notation

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

$$\begin{aligned} y_{1,t} &= a_{11}\ y_{1,t-1} + a_{12}\ y_{2,t-1} + c_1 + u_{1,t} \\ y_{2,t} &= a_{21}\ y_{1,t-1} + a_{22}\ y_{2,t-1} + c_2 + u_{2,t} \end{aligned}$$

- The variance-covariance matrix of residuals $u_{1,t}$ and $u_{2,t}$ is:

$$E(u_t u_t') = \Sigma = \begin{bmatrix} var(u_{1t}) & cov(u_{1t}, u_{2t}) \\ cov(u_{2t}, u_{1t}) & var(u_{2t}) \end{bmatrix} = \begin{bmatrix} \Sigma_{u_1} & \Sigma_{u_1 u_2} \\ \Sigma_{u_2 u_1} & \Sigma_{u_1} \end{bmatrix}$$

Example

- 2^{nd} specification: Matrix notation. Note. This is how we mainly specify a VAR in statistical softwares (i.e R, Matlab, ...)

$$\underbrace{\begin{bmatrix} y_{1,1} & y_{2,1} \\ y_{1,2} & y_{2,2} \\ \vdots & \vdots \\ y_{1,T-1} & y_{2,T-1} \\ y_{1,T} & y_{2,T} \end{bmatrix}}_{(T\times K)} = \underbrace{\begin{bmatrix} y_{1,0} & y_{2,0} & 1 \\ y_{1,1} & y_{2,1} & 1 \\ \vdots & \vdots & \vdots \\ y_{1,T-2} & y_{2,T-2} & 1 \\ y_{1,T-1} & y_{2,T-1} & 1 \end{bmatrix}}_{T\times(K+1)} \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ c_1 & c_2 \end{bmatrix}}_{(K+1)\times K} + \underbrace{\begin{bmatrix} u_{1,1} & u_{2,1} \\ u_{1,2} & u_{2,2} \\ \vdots & \vdots \\ u_{1,T-1} & u_{2,T-1} \\ u_{1,T} & u_{2,T} \end{bmatrix}}_{(T\times K)}$$

$$Y = XA + U$$

Companion matrix of a VAR(p) process

- A K-dimensional VAR with a lag of order p can be written as a Kp-dimensional VAR(1):

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + c + u_t \quad , \quad \text{with} \quad u_t \sim \mathcal{N}(0, \Sigma)$$

- In companion form:

$$\underset{(Kp \times 1)}{\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p+1} \end{bmatrix}} = \underset{(Kp \times Kp)}{\begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p \\ I_K & 0 & \ldots & 0 & 0 \\ 0 & I_K & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & I_K & 0 \end{bmatrix}} \underset{(Kp \times 1)}{\begin{bmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ \vdots \\ y_{t-p} \end{bmatrix}} + \underset{(Kp \times 1)}{\begin{bmatrix} c \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} + \underset{(Kp \times 1)}{\begin{bmatrix} u_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$

Companion matrix of a VAR(p) process

- Hence, the VAR(p) is written as first order system:

$$Y_t = \mathbf{A} Y_{t-1} + \mathbf{c} + U_t$$

with a variance covariance matrix of the Kp-dimensional vector of VAR residuals:

$$\mathbf{\Sigma} = E(U_t U_t') = \underset{(Kp \times Kp)}{\begin{bmatrix} \Sigma & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}$$

where $\Sigma = E(u_t u_t')$

VAR estimation

- Consider a K-dimensional VAR(p):

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + c + u_t$$

- and use the 2^{nd} specification of the VAR (matrix notation - compact form):

$$Y = XA + U$$

where:

$$\begin{aligned}
Y &= [y_{1t}, y_{2t}, \ldots, y_{kt}] : && T \times K \\
X &= [y_{t-1}, y_{t-2}, \ldots, y_{t-p}, \mathit{1}] : && T \times (Kp + 1) \\
A &= [A'_1, A'_2, \ldots, A'_p, c']' : && (Kp + 1) \times K \\
U &= [u_{1t}, u_{2t}, \ldots, u_{kt}] : && T \times K
\end{aligned}$$

VAR estimation - Ordinary least square

- The VAR can be estimated by ordinary least square (OLS) which results in efficient estimator of the VAR coefficients:

$$\hat{A}^{LS} = [\hat{A}_1', \hat{A}_2', \ldots, \hat{A}_p', \hat{c}']' = (X'X)^{-1}X'Y$$

  which is equivalent to estimating the VAR equation by equation separately, using OLS.

- The (consistent) OLS estimator of the residuals variance-covariance matrix is:

$$\widehat{\Sigma}^{LS} = \frac{\widehat{U}'\widehat{U}}{T-(Kp+1)} \quad , \quad \text{where} \quad \widehat{U} = Y - X\hat{\mathrm{A}}$$

- The variance-covariance matrix of the VAR parameters in $\hat{\mathrm{A}}$ is:

$$var(\hat{\mathrm{A}})^{LS} = \widehat{\Sigma}^{LS} \otimes (X'X)^{-1}$$

VAR estimation - Maximum likelihood

- Alternatively, the VAR parameters can be estimated by Maximum Likelihood.
- Under the assumption that u_t are independent and identically distributed, the multivariate gaussian density of y_t is:

$$\frac{1}{(2\pi)^{K/2} det(\Sigma)^{1/2}} \; exp\Big(- \frac{1}{2} u_t' \Sigma^{-1} u_t \Big)$$

with a corresponding log-density:

$$-\frac{K}{2} log(2\pi) - \frac{1}{2} log\Big[det(\Sigma) \Big] - \frac{1}{2} u_t' \Sigma^{-1} u_t$$

and the log-likelihood:

$$-\frac{KT}{2} log(2\pi) - \frac{T}{2} log\Big[det(\Sigma) \Big] - \frac{1}{2} \sum_{t=1}^{T} (u_t' \Sigma^{-1} u_t)$$

VAR estimation - Maximum likelihood

- The ML estimators of the VAR coefficients and of the variance-covariance matrix of the innovations can be obtained by maximizing the log-likelihood.
- Generally:

$$\hat{\text{A}}^{ML} \equiv \hat{\text{A}}^{LS}$$

- The ML estimator of the residuals variance-covariance matrix is:

$$\widehat{\Sigma}^{ML} = \frac{\widehat{U}'\widehat{U}}{T}$$

Lag length selection

- The choice of the lag order (p) is important since it affects the number of parameters in a VAR.
- In relatively small sample size, this could potentially affect the precision of the estimates.
- For example, with $K = 5$ variables and $p = 1$:

$$\text{N. parameters} = (Kp + 1) \times K = (5 \times 1 + 1) \times 5 = 30$$

- with $K = 5$ variables and $p = 2$:

$$\text{N. parameters} = (Kp + 1) \times K = (5 \times 2 + 1) \times 5 = 55$$

- with $K = 5$ variables and $p = 4$:

$$\text{N. parameters} = (Kp + 1) \times K = (5 \times 4 + 1) \times 5 = 105$$

- The number of parameters in a VAR increases dramatically as the lag order rises, by a rate of $K^2 \times r$, with r being the difference between the two lags.

Lag length selection

- There are statistical criteria to select the lag order. These are based on selecting the lag order that minimizes:
  - AIC: Akaike information criterion

$$AIC(p) = log\left(det(\hat{\Sigma}_p)\right) + \frac{2(K^2p + K)}{T}$$

  - BIC: Bayesian information criterion

$$BIC(p) = log\left(det(\hat{\Sigma}_p)\right) + \frac{(K^2p + K)log(T)}{T}$$

  - HQ: Hannan-Quinn information criterion

$$HQ(p) = log\left(det(\hat{\Sigma}_p)\right) + \frac{2(K^2p + K)log(log(T))}{T}$$

- Another approach is based on sequential testing. The models with different lag orders are compared based on a Wald or on Likelihood ratio (LR) test.

Structural VAR (SVAR)

- Consider a K-dimensional VAR(p):

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + c + u_t \quad , \quad \text{with} \quad u_t \sim \mathcal{N}(0, \Sigma)$$

- The residuals, u_t, are correlated: it is not possible to attach an economic interpretation $\rightarrow$ **reduced form VAR**.
  The disturbances need to be uncorrelated (orthogonal).

- Consider the following structural VAR(p) representation:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + c + B_0 \varepsilon_t \quad , \quad \text{with} \quad \varepsilon_t \sim \mathcal{N}(0, \Omega)$$

- The variance-covariance matrix of the structural shocks, Ω, is generally normalized such that:

$$E(\varepsilon_t \varepsilon_t') = \Omega = I_K$$

Structural VAR (SVAR)

- The relationship between reduced form innovations and structural shocks is:

$$u_t = B_0 \varepsilon_t$$

$$\underbrace{E(u_t u_t')}_{\Sigma} = E(B_0 \varepsilon_t \varepsilon_t' B_0') = B_0 \underbrace{E(\varepsilon_t \varepsilon_t')}_{I_K} B_0' = B_0 B_0'$$

$$\Sigma = B_0 B_0'$$

- Several ways of estimating B_0 (identification strategy) that will be discussed later. For the time being, let us assume that we have an estimate of B_0.
- Once retrieved estimates of B_0, it is possible to conduct **structural dynamic analysis**.

Road Map

- Teaching materials: notes, R codes including:
  - main.R (script containing the main code).
  - packages (to be used for installing the required R packages).
  - functions (folder containing *ad-hoc* functions).
  - data used in the illustrative example.
- Illustrative example: identification of oil price shock and VAR tools.

main.R

```
#=========================================================================================
#                                main.R
#=========================================================================================
# This script replicates the example in Kilian and Lutkepohl (2017), Structural Vector Autoregressive
# Analysis, Cambridge University Press, Chapter 9 (page 239-240).
#
# The exercise focuses on the identification of an oil price shock and its impact on inflation and
# real GDP.
#=========================================================================================

# 1. Setting work directory, installing packages and importing functions

# 1.1 Setting work directory
wd <- 'C:\\namefolder\\SER\\' # or alternatively one can set the wd manually

# Setting work directory
setwd(wd)

# 1.2 Updating packages
source(paste0(wd, 'packages\\packages.R'))

# 1.3 Updating functions
file.sources <- list.files(paste0(wd, '\\functions') , pattern="*.R$", full.names=TRUE, ignore.case=FALSE)
sapply(file.sources, source, .GlobalEnv)
#-----------------------------------------------------------------------------------------

# 2. Import data and visualizing series
# Time series used in the exercise:
#
# RPOIL = percent changes in the real WTI price of crude oil
# INFL  = U.S. GDP deflator inflation rate
# GDP   = U.S. real GDP growth

# 2.1. Import data and labels
DATAFULL <- read.csv(paste0(wd, 'data\\dataoil1.csv'), header = TRUE)

# ... (continue)
```

See “main.R”

Packages

```
#=====================================================================
#         packages.R
#=====================================================================
# This script uploads packages.
# Notes. Uncomment "install.packages" if the package has not been
# installed, yet.
#=====================================================================

# install.packages("Matrix")
library(Matrix)

# install.packages("expm")
library(expm)

# install.packages("ggplot2")
library(ggplot2)

# install.packages("gridExtra")
library(gridExtra)
```

Functions (example)

```
#==========================================================================================
#                    getlag.R
#==========================================================================================
# This function constructs the matrix containing the lagged endogenous
# variables plus the intercept
#==========================================================================================

getlag <- function(DATAMAT,plag,CONSTANT) {

if (is.matrix(DATAMAT) == FALSE) {
  DATAMAT <- matrix(DATAMAT)
}

kvar <- ncol(DATAMAT)
TOBS <- nrow(DATAMAT)
TESS <- TOBS - plag

X <- matrix(0, nrow = TESS, ncol = 0)

# Lagged endogenous variables
for (tt in 1 : plag) {

  X <- cbind(X , DATAMAT[((1+plag)-tt) : (TOBS-tt) , ])

}

if (!is.null(colnames(DATAMAT))) {
  colnames(X) <- paste0(rep(colnames(DATAMAT) , plag), '_l', rep(1:plag, each=kvar))
} else {
  colnames(X) <- paste0(rep(paste0('y', 1:kvar), plag), '_l', rep(1:plag, each=kvar))
}

# Constant
if (CONSTANT == 1) {

  X <- cbind(X , rep(1 , TESS))
  colnames(X)[ncol(X)] <- 'constant'

} else if(CONSTANT == 0) {
  X <- X

}

return(X)
}
```

Data

- Data are from the empirical applications in Kilian and Lutkepohl (2017), Chapter 9, Figure 9.1.
- Quarterly series on real West Texas Intermediate (WTI) price of crude oil, U.S. GDP deflator inflation rate and U.S. real GDP.
- Estimation sample: 1973Q1 – 2013Q2.
- Data on WTI oil price are from Economagic, while U.S. GDP deflator and U.S. real GDP are downloaded from the Federal Reserve Bank of St. Louis (FRED) Database.

Illustrative example

Identification through Cholesky decomposition

Short-run (zero exclusions) restrictions

- See Kilian and Lütkepohl (2017, Chapters 8 and 9).
- Consider a reduced form VAR(p):

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + c + u_t \quad , \quad u_t \sim \mathcal{N}(0, \Sigma)$$

- and the corresponding structural form VAR(p):

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + c + B_0 \varepsilon_t \quad , \quad \varepsilon_t \sim \mathcal{N}(0, \underbrace{\Omega}_{I_K})$$

- The relationship between the reduced form innovations and the structural disturbances holds:

$$u_t = B_0 \varepsilon_t$$
$$\Sigma = E(u_t u_t') = B_0 \underbrace{E(\varepsilon_t \varepsilon_t')}_{\Omega = I_K} B_0'$$

Short-run (zero exclusions) restrictions - recursive scheme

- Suppose we have K=3 endogenous variables:

$$\Sigma = B_0 B_0'$$

$$\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ & s_{22} & s_{23} \\ & & s_{33} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix}$$

- This relationship produces a system of $\frac{K(K+1)}{2} = 6$ equations with $K^2 = 9$ unknown (free) parameters: $[b_{11}, b_{12}, b_{13}, b_{21}, b_{22}, b_{23}, b_{31}, b_{32}, b_{33}]$.

- To exactly identify the system of equations we need to impose $\frac{K(K-1)}{2} = 3$ zero restrictions.

Short-run (zero exclusions) restrictions - recursive scheme

- One popular way to impose zero contemporaneous restrictions is through Cholesky factorization (decomposition).
- Any singular positive definite matrix (Σ) can be decomposed into:

$$\Sigma = PP'$$

  where P is a lower triangular matrix.
- Since:

$$\Sigma = B_0 B_0'$$

- It follows that:

$$B_0 = P$$

Illustrative example

- Kilian and Lütkepohl (2017, Chapter 9, page 239 – 240)
- See “main.R” script.
- Suppose we have $K = 3$ endogenous variables: $\Delta rpoil_t$, Δp_t and Δgdp_t.
- $\Delta rpoil_t$ is the percent changes in the real WTI price of crude oil, Δp_t is the U.S. GDP deflator inflation rate and Δgdp_t is the U.S. real GDP growth.
- VAR(4):

$$y_t = \sum_{i=1}^{4} A_i y_{t-i} + c + u_t \quad , \quad u_t \sim \mathcal{N}(0, \Sigma)$$

$$\text{Since that } VAR(\varepsilon_t) = \Omega = I_K \quad \Rightarrow \quad E(u_t u_t') = \Sigma = B_0 B_0'$$

- The structural shock of interest is an oil price shock ($\varepsilon_{rpoil,t}$).

Endogenous variables

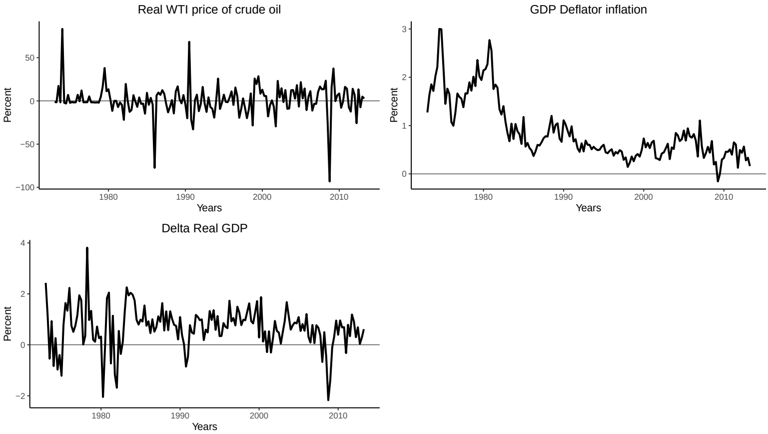

Figure 2: Percent changes in the real WTI price of crude oil, U.S. GDP deflator inflation rate and U.S real GDP growth, 1973Q1 – 2013Q2.

Cholesky decomposition

- Impose a recursive ordering of the variables and compute the Cholesky decomposition of the reduced form residuals covariance matrix, Σ:

$$\begin{bmatrix} u_{rpoil,t} \\ u_{p,t} \\ u_{gdp,t} \end{bmatrix} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}}_{B_0 = P = chol(\Sigma)} \begin{bmatrix} \varepsilon_{rpoil,t} \\ \varepsilon_{p,t} \\ \varepsilon_{gdp,t} \end{bmatrix}$$

Identification strategy

- The impulse responses of the three variables to the oil price shock on impact ($h = 0$) are:

$$\begin{bmatrix} \frac{\delta rpoil_{t+h}}{\delta \varepsilon_{rpoil,t}} \\ \frac{\delta p_{t+h}}{\delta \varepsilon_{rpoil,t}} \\ \frac{\delta gdp_{t+h}}{\delta \varepsilon_{rpoil,t}} \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

- The real WTI price of crude oil is not affected contemporaneously by disturbances to inflation and GDP. The real price of oil respond contemporaneously to the oil price shock (b_{11}).
- GDP and Inflation react contemporaneously to oil price shocks (b_{21}, b_{31}).
- The shock to the real price of oil is interpreted as oil price shock.
- No economic interpretation of the two other shocks.

Impulse Response analysis

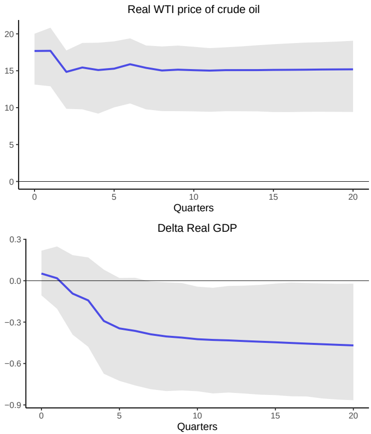

Figure 3: Impulse response function of the variables to oil price shock.

Forecast Error Variance decomposition

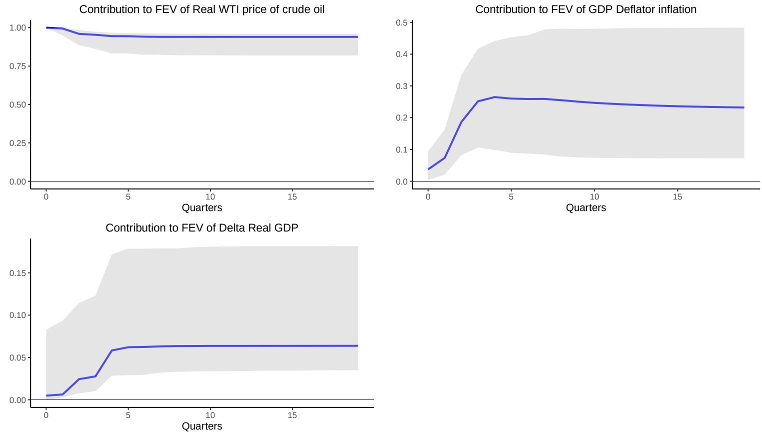

Figure 4: Contribution of oil price shock to the FEV of the variables.

Historical decomposition

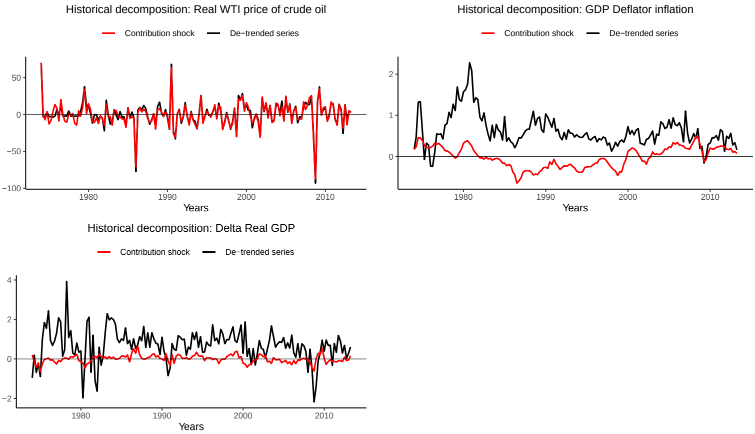

Figure 5: Historical contribution of oil price shock to the de-trended variables.

References: R packages

- Douglas Bates, Maechler, Martin, and Davis Timothy A., 2021. "Matrix", Version 1.3-3. URL: https://cran.r-project.org/web/packages/Matrix/index.html
- Vincent Goulet, Dutang, Christophe, Maechler, Martin, Firth, David, Shapira, Marina, and Stadelmann, Michael, 2021. "expm", Version 0.999-6. URL: https://cran.r-project.org/web/packages/expm/index.html
- Hadley Wickham, Chang, Winston, Henry, Lionel, Pedersen, Thomas L., Takahashi, Kohske, Wilke, Claus, Woo, Kara, Yutani, Hiroaki, Dunnington, Dewey, RStudio, 2020. "ggplot2", Version: 3.3.3. URL: https://cran.r-project.org/web/packages/ggplot2/index.html
- Baptiste Auguie, Antonov, Anton, 2017. "gridExtra", Version: 2.3. URL: https://cran.r-project.org/web/packages/gridExtra/index.html