The VAR Toolbox Handbook[☆]

Ambrogio Cesa-Bianchi[†]

March 29, 2023

Abstract

This paper describes a collection of Matlab routines to perform VAR analysis. The VAR toolbox allows to estimate reduced-form VARs and to identify structural shocks under a number of different identification techniques. Impulse responses, forecast error variance decompositions, and historical decompositions are computed according to the chosen identification option. The toolbox includes practical examples and replications of well-known studies in the VAR literature.

[☆]The views expressed in this paper are those of the author and do not necessarily represent the views of the Bank of England.

[†]Bank of England, CEPR, and CFM.
Email: ambrogio.cesa-bianchi@bankofengland.co.uk.
Web: https://sites.google.com/site/ambropo

Contents

1 VAR Toolbox: High level description

The VAR Toolbox is a collection of Matlab routines to perform Vector autoregressive (VAR) analysis. Initially developed by Sims (1980) more than 40 years ago, as of today VAR models are still one of the most important tools in a macroeconometrician's toolkit. As famously put it by Stock and Watson (2001), VARs provide a coherent and credible approach for macroeconometricians to do their job, which consists of four things: (i) describing the dynamic relations in macroeconomic and financial data, (ii) making forecasts, (iii) making inference about the true (and unobserved) structure of the macroeconomy, and (iv) advising policymakers.

The VAR Toolbox provides an intelligible and accessible way into the workings of VARs step by step. In this regard, it differs from similar (and more efficient) toolboxes, such as the BEAR toolbox of Dieppe, van Roye and Legrand (2016) and the hitchhiker guide to empirical macro models of Canova and Ferroni (2020).[1] Estimation is performed with Ordinary Least Squares (OLS) and confidence intervals are obtained with bootstrapping methods. The toolbox includes the most commonly used identification options, and allows the computation of impulse responses, forecast error variance decompositions, and historical decompositions.

The objective of this paper is to explain the functioning of the VAR Toolbox and the workings of VARs by means of simple examples – the idea being that is much easier to learn by doing rather than reading a technical manual first, and then go to the computer. As a result, many of the functions included in the VAR Toolbox are not covered here, nor is a full description of the output of each function. Moreover, to simplify the intuition and exposition, most of the material covered here is treated in a pretty informal way. In other words, the VAR Toolbox handbook should not be thought of as a substitute for more formal time series econometrics textbooks, but rather as a complement to them.

[1] Other toolboxes covering similar topics are the Econometric toolbox of James LaSage, the Dynare project of Adjemian, Bastani, Juillard, Karam e, Maih, Mihoubi, Perendia, Pfeifer, Ratto and Villemot (2011), the Global VAR toolbox of Vanessa Smith, the mixed frequency VAR toolbox of Brave, Butters and Kelley (2020), the Bayesian Local Projection of Miranda-Agrippino and Ricco.

Acknowledgements: the origin of the VAR Toolbox traces back to the early days of my PhD, when I was trying to understand the mechanics of VARs by replicating existing papers in the literature. In my personal experience, replication is key: it is only when coding up something that you really get the grips with it. The VAR Toolbox would not exist if it were not for James LeSage's Econometrics Toolbox, which helped greatly with my understanding of VARs and econometrics more in general. The main function for the estimation of reduced form VARs in the VAR Toolbox is still a slightly modified version of LeSage's original function. Relative to Le Sage's Toolbox, the VAR Toolbox is much narrower in scope, its main focus being the estimation and structural identification of VAR models. The VAR Toolbox has then developed over the years, with the help of many users, colleagues, and co-authors who made useful suggestions and spotted typos or bugs with the codes. I'm grateful to all of you.

Disclaimer: All files available in the VAR toolbox are for education and/or research purposes only. I take no responsibility and/or liability for how you choose to use any of the source code available here. While the codes in the VAR Toolbox have been tested extensively, and despite every effort has been made to ensure that they are error free, some of them may still have bugs or errors. If you find any, please email me at `ambrogio.cesabianchi@gmail.com` or open an issue in Github `https://github.com/ambropo/VAR-Toolbox`. Whenever the software is used, I would appreciate acknowledgment by citation of this working paper.

2 Getting started

No installation is required. Simply fork/download the latest version of the toolbox from `https://github.com/ambropo/VAR-Toolbox` to a specific folder on the hard drive, e.g. `/VARToolbox/v3dot0`. The only required step to get the VAR Toolbox working is to add the folder `v3dot0` (including all subfolders) to the Matlab path. To avoid clashes with other function it is recommendable to add and remove the Toolbox with the following commands at both the beginning and the end of your scripts, e.g.:

```
addpath(genpath('/VARToolbox/v3dot0'))
...
rmpath(genpath('/VARToolbox/v3dot0'))
```

The codes are grouped in six categories (and respective subfolders within the master folder `/VARToolbox/v3dot0`):

- `VAR`: codes for VAR analysis, e.g. estimation, identification, computation of the impulse response functions, forecast error variance decompositions, historical decompositions, etc.
- `Stats`: codes for the calculation of commonly used descriptive statistics, e.g. moving-window averages or sums, pairwise correlations, etc.
- `Utils`: codes that allow the smooth functioning of the Toolbox, e.g. functions to vectorize matrices or compute the number of rows of a matrix.
- `Auxiliary`: codes that I borrowed from other public sources. Each m-file has a reference to the original source.
- `Figure`: codes for plotting high quality figures, particularly thought for time series data, e.g. functions to control dates on the horizontal axis, appearance of the legends, plot charts with shaded error bands, etc.
- `ExportFig`: codes developed by Yair Altman for exporting high quality figures, available at https://github.com/altmany/export_fig. To enable this option, the VAR Toolbox requires Ghostscript installed on your computer (freely available at www.ghostscript.com).

Two additional folders include the codes for all examples and replications presented in this paper and its accompanying slide deck:

- `Primer`: Codes for all the examples used in this paper.
- `Replic`: Codes for the replication of a few well-known VAR studies.

The next sections describe how to load the data, estimate a VAR, identify the shocks of interest, and analyze how these shocks transmit through the system, as well as how important they are in driving variation in the endogenous variables. The sections closely follow the code `VARToolbox_Primer.m` in the folder `../Primer/`.

At the time of writing, the latest version of the toolbox is 3.0. The VAR Toolbox 3.0 has been tested with Matlab R2021B on a Macbook Pro machine.

3 Prelims: Loading and Preparing the Data

The data for the examples in this paper is stored in a spreadsheet (`Simple_Data.xlsx`) stored in `../Primer/data/`. The Data Appendix reports the source of each series used.

The file `Simple_Data.xlsx` contains US macro and financial data at quarterly frequency, namely the CPI Index, real GDP, the unemployment rate, the VIX Index, and the yield on the 1-year Treasury Bill. The sample period is 1989:Q1 to 2019:Q4, so that the number of observations for each time series is $T = 124$.

The code below shows a general way of loading the data and managing it in a way that is consistent with the functioning of the VAR Toolbox. Specifically, the code reads from the spreadsheet `Simple_Data.xlsx` and stores each time series into the structure `DATA` as a separate variable. The convention of the VAR Toolbox is that time series are stored in column-vectors, so that the number of rows corresponds to the numbers of observations (denoted by T) for a given time series.

The VAR Toolbox can deal with dates in both string and numeric format. Dates in string format follow the usual convention that quarters (months) are denoted with Q (M); for example, `'2000Q1'` (`'2000M1'`) corresponds to the first quarter (month) of the year 2000. Dates in numeric format, as in Canova and Ferroni (2020), follow the convention that the first quarter (month) of the year corresponds to the integer of that year; for example, 2000Q1 (or 2000M1 for monthly data) corresponds to `2000.00`. In the above example, the dates are read in string format from the spreadsheet `Simple_Data.xlsx` and stored in the cell array `dates`. The user can switch from one convention to the other with the functions `Date2Num`, which converts a vector of dates from cell to double arrays, and `Date2Cell`, which converts a vector of dates from double to cell arrays.

```
% Load data from US macro data set
[xlsdata, xlstext] = xlsread('data/Simple_Data.xlsx','Sheet1');
dates = xlstext(3:end,1);          % vector of dates in string format
datesnum = Date2Num(dates);        % vector of dates in numeric format
vnames_long = xlstext(1,2:end);  % full variable names
vnames = xlstext(2,2:end);        % variable mnemonic
nvar = length(vnames);             % number of variables in
spreadsheet
data   = Num2NaN(xlsdata);          % matrix of data in spreadsheet
% Store variables in the structure DATA
for ii=1:length(vnames)
   DATA.(vnames{ii}) = data(:,ii);
end
% Observations
nobs = size(data,1);
```

As it is the case in many empirical applications, the raw data needs to be transformed to be used in VAR analysis. The VAR Toolbox has some built-in functions to perform the most commonly used data treatments. The example below shows how to compute the growth rate of real GDP and the CPI Index, and the first difference of the 1-year Treasury Bill yield. The new variables are added to the structure `DATA`.

```
% Select variables to treat
tempnames = {'gdp','cpi','i1yr'};          % variable mnemonics
temptreat = {'logdiff','logdiff','diff'}; % type of transformation
% Treat and add to DATA structure
for ii=1:length(tempnames)
   aux = {['d' tempnames{ii}]};
   DATA.(aux{1}) = XoX(DATA.(tempnames{ii}),1,temptreat{ii});
end
delete temp*
```

The example used throughout the paper is based on an (overly) simplistic VAR where the only $k = 2$ endogenous variables are the quarterly growth rate of US GDP (as computed above and denoted by y_t) and the 1-year yield on the US Treasury bill (denoted by r_t). While such a simple VAR cannot realistically describe the complex interactions of the US economy, it is a useful device to understand the functioning of the codes in the VAR Toolbox codes and, importantly, the mechanics of VAR models more in general.[2]

The VAR Toolbox follows the convention that each endogenous variable is stored in a column vector of a $T \times k$ matrix `X`, as follows:

$$\texttt{X} = \begin{bmatrix} y_1 & r_1 \\ y_2 & r_2 \\ ... & ... \\ y_T & r_T \end{bmatrix} \tag{1}$$

so that the matrix `X` has $T = 124$ rows (i.e. the number of observations) and $k = 2$ columns (i.e. the number of variables). The code below shows a general way to construct such `X` matrix in Matlab.

```
% Select the list of endogenous variables...
Xvnames = {'dgdp','i1yr'};
% ... and corresponding labels to be used in plots
Xvnames_long = {'Real GDP Growth','1-year Int. Rate'};
% Number of endo variables
Xnvar = length(Xvnames);
% Create matrix X of variables to be used in the VAR
X = nan(nobs,Xnvar);
for ii=1:Xnvar
    X(:,ii) = DATA.(Xvnames{ii});
end
```

[2] The VAR Toolbox also includes more realistic examples based on replications of existing papers. The replication codes, which are not discussed in this paper, can be found in the folder `/VARToolbox/v3dot0/Replic/`. The replication examples are described in the slide deck that accompanies this handbook.

The VAR Toolbox includes some functions that allow to plot time series quickly and export them as high-quality PDFs, so that they can be used directly in research papers. The code shows how to plot the two time series in `X`.

```
% Open a figure of the desired size and plot the selected variables
FigSize(26,8)
for ii=1:Xnvar
   subplot(1,2,ii)
   H(ii) = plot(X(:,ii),'LineWidth',3,'Color',cmap(1));
   title(Xvnames_long(ii));
   DatesPlot(datesnum(1),nobs,6,'q') % Set the x-axis labels
   grid on;
end
% Save figure
SaveFigure('graphics/BIV_DATA',1)
clf('reset')
```

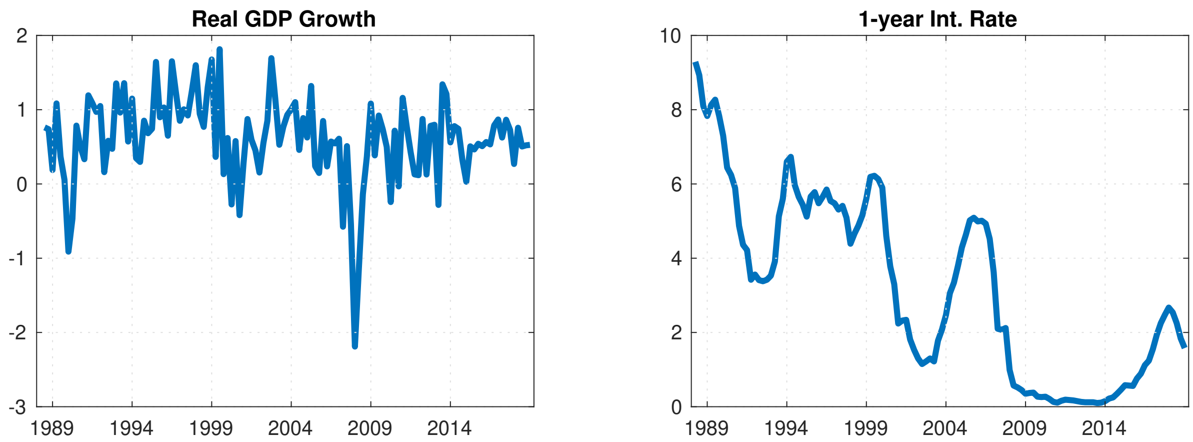
Some useful functions used in the code above are:

- `FigSize.m`: allows the user to choose the size and the proportions of the window for plotting the figure. This is particularly useful when creating figures with many panels.
- `DatesPlot.m`: Adds the dates (in numeric format) to the horizontal axis of a chart (at monthly, quarterly, or annual frequency) using a specified number of ticks (6 in the above example).
- `SaveFigure.m`: saves the chart in the selected format (pdf, jpg, eps). The function allows the user to save the figure at high quality standard using the `export_fig.m` function developed by Yair Altman.[3]

Figure 1 reports the behavior the growth rate of real GDP and the interest rate on the US 1-year Treasury bill over the 1989:Q1 to 2019:Q4 sample period.

[3] Original code is available at https://github.com/altmany/export_fig.

Figure 1 ENDOGENOUS VARIABLES IN THE SIMPLE VAR



NOTE. Growth rate of real GDP US and the 1-year Treasury bill yield. Percentage points. Sample period: 1989:Q1 to 2019:Q4.

4 A Simple VAR Model

The main idea of this paper is to employ a simple example – containing the minimal possible number of ingredients – to describe the workings of VAR models. A minimal example makes the exposition easier to follow, the algebra trivial, and allows to double check in a straightforward way the calculations performed by Matlab . These advantages, however, come at the cost of some unrealistic assumptions, which affect the economic interpretation of the results. In most cases, the examples that follow are too simple and parsimonious to credibly identify structural shocks (such as monetary policy shocks, or demand and supply shocks) and approximate the true structure of the economy. The VAR Toolbox also includes more realistic examples based on replications of existing papers. The replication codes, which are not discussed in this paper, can be found in the folder `../Replic/`.

Consider a simple bivariate VAR(1), i.e. a VAR where the number of endogenous variables is $k = 2$ and the number of lags is $p = 1$, with a constant term. Let the two endogenous variables be output growth (y_t) and the interest rate on the 1-year Treasury Bill (r_t). The reduced-form representation of the bivariate VAR(1) can be written as:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{rt} \end{bmatrix} \tag{2}$$

or:

$$\begin{aligned} y_t &= c_y + \phi_{11} y_{t-1} + \phi_{12} r_{t-1} + u_{yt}, \\ r_t &= c_r + \phi_{21} y_{t-1} + \phi_{22} r_{t-1} + u_{rt}, \end{aligned} \tag{3}$$

In matrix form, the VAR in (2) can be written more compactly as

$$x_t = c + \Phi_1 x_{t-1} + u_t, \tag{4}$$

where x_t is $2 \times T$ matrix collecting the two endogenous variables;[4] c is a 2×1 vector of constants; Φ_1 is a 2×2 matrix of autoregressive coefficients; and u_t is a $2 \times T$ vector of serially uncorrelated innovations, generally referred to as **reduced-form residuals**. Typically, the reduced-form residuals are correlated among themselves. Their covariance matrix can be written as:

$$\mathbb{V}(u_t) \equiv \Sigma_u = \begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ \sigma_{yr}^2 & \sigma_r^2 \end{bmatrix}. \tag{5}$$

The covariance matrix Σ_u is 2×2 symmetric matrix. Its diagonal elements are the variances of the estimated reduced-form innovations, σ_y^2 and σ_r^2; and the identical off-diagonal elements are instead equal to the covariance between the estimated reduced-form residuals, σ_{yr}^2. The covariance between the estimated reduced-form residuals plays an important role in VARs because it collects the information on the contemporaneous relation among the variables in the system (after controlling for their persistence) – which, as we shall see below, is a key element of the identification problem in VARs.

The estimation of a VAR model can be done with a simple line of code, using the `VARmodel.m` function. There are three crucial inputs to this function: (i) a matrix

[4] Note that, following the notation of the previous section, x_t is the transpose of `X`.

including the endogenous variables, `X`; (ii) an integer specifying the number of lags to use, `nlag`; and an integer specifying whether a constant or a trend should be included, `det`. The code below shows how to estimate the simple VAR model in (2).

```
% Make a common sample by removing NaNs
[X, fo, lo] = CommonSample(X);
% Set the deterministic variable in the VAR (1=constant, 2=trend)
det = 1;
% Set number of lags
nlags = 1;
% Estimate VAR by OLS
[VAR, VARopt] = VARmodel(X,nlags,det);
```

The results of the VAR estimation are stored in the structures `VAR` and `VARopt`. These are two crucial objects in the VAR Toolbox. Not only they collect all the relevant information of the estimated VAR, but they are also contain various options for identification (as shown below) and are updated with new information after each step of the VAR analysis.[5]

The structure `VAR` includes all the inputs to the `VARmodel.m` function – such as the matrix of endogenous variables (`VAR.ENDO`), the chosen number of lags (`VAR.nlags`), the number of endogenous variables (`VAR.nvar`), etc). Crucially, the structure `VAR` also includes the estimation output. These can be seen at screen by double-clicking on the structure `VAR` in the Matlab workspace or by printing its output in the command window with the following command:

```
>> disp(VAR)
        ENDO: [123x2 double]
        nlag: 1
        const: 1
        EXOG: []
        nobs: 122
        nvar: 2
```

[5]For those familiar with Dynare, the structures `VAR` and `VARopt` are similar in spirit to the `_m` and `_oo` structures.

```
nvar_ex: 0
nlag_ex: 0
ncoeff: 2
ntotcoeff: 3
eq1: [1x1 struct]
eq2: [1x1 struct]
Ft: [3x2 double]
F: [2x3 double]
sigma: [2x2 double]
resid: [122x2 double]
X: [122x3 double]
Y: [122x2 double]
Fcomp: [2x2 double]
maxEig: 0.9559
B: []
b: []
PSI: []
Fp: []
IV: []
```

A few of the elements of the structure `VAR` are worth describing in detail

- **Estimated coefficients**. The matrix `VAR.F` collects all estimated coefficients following the notation in (2), so that `VAR.F` $= [c\ \Phi]$. For a VAR with 1 lag and 2 endogenous variables plus a constant, this means that `VAR.F` is a $2 \times (1 + 1 \times 2)$ matrix, as shown by the following command:

```
>> disp(VAR.F)
     0.3630    0.3788    0.0041
    -0.0729    0.2607    0.9541
```

- **VAR residuals** The matrix `VAR.resid` collects the VAR reduced-form residuals, as defined by (2). In line with the convention in equation (1) the residuals are stored as column vectors so that `VAR.resid` $= u_t'$. That is, for a bivariate VAR with 1 lag, `VAR.resid` is a $2 \times (T-1)$ matrix (as one observation gets lost when computing the lags of x_t).

- **Reduced-form covariance matrix** The matrix `VAR.sigma` collects the covariance matrix of the VAR reduced-form residuals defined by (5). The convention is such that `VAR.sigma` $=\Sigma_u$. For a VAR with 2 endogenous variables, the covariance matrix `VAR.sigma` is a symmetric 2×2 matrix:

```
>> disp(VAR.sigma)
                0.2891    0.0782
                0.0782    0.1473
```

- **Companion matrix** The matrix `VAR.Fcomp` includes the *companion matrix*. The companion matrix allows rewriting VARs with lags greater than 1 as VAR(1), see Appendix B for details. Trivially, in the case of a VAR(1) the companion matrix `VAR.Fcomp` is identical to `VAR.F` after dropping deterministic coefficients, such as the constant or trend. As it will become clear below, the companion matrix plays a crucial role for many aspects of VAR analysis, such as the computation of impulse responses, the identification of structural shocks, etc.

- **Equation-by-equation estimation output**. The structures `VAR.eq1` and `VAR.eq2` include the OLS equation-by-equation estimation results.

Note that the VAR structure includes a few empty objects (e.g. `VAR.B`). This is because, for the moment, the code has only estimated the reduced-form VAR, and has not yet identified the structural shocks or computed impulse response functions. This will be the object of the next section.

The structure `VARopt` includes a few auxiliary variables that are created automatically by the `VARmodel.m` function and will be needed below for the calculation of impulse responses, variance decompositions, etc. As above, the variables stored in `VARopt` can be seen by executing:

```
>> disp(VARopt)
        vnames: []
        vnames_ex: []
        snames: []
        nsteps: 40
        impact: 0
        shut: 0
```

```
ident: 'short'
recurs: 'wold'
ndraws: 1000
mult: 10
pctg: 95
method: 'bs'
sr_hor: 1
sr_rot: 500
sr_draw: 100000
sr_mod: 1
pick: 0
quality: 1
suptitle: 0
datesnum: []
datestxt: []
datestype: 1
firstdate: []
frequency: 'q'
figname: []
FigSize: [26 24]
```

These variables include the number of steps for impulse response functions and variance decompositions (`nsteps`), the labels of the endogenous variables for plots (`vnames`), the confidence levels for the computation of confidence intervals (`pctg`), etc. While some variables are automatically created by the `VARmodel` function (i.e. where a default option is possible), some other variables need to be inputted by the user. For example, the code below manually adds to the `VARopt` structure the names of the VAR endogenous variables:

```
% Print at screen the outputs of the VARmodel estimation
disp(VAR)
disp(VARopt)
% Update the VARopt structure with additional details
VARopt.vnames = Xvnames_long;
```

Finally, the VAR Toolbox has a built-in function to print at screen some of the most important estimation results in an easy and quick way:

```
% Print at screen VAR coefficients and create table
[TABLE, beta] = VARprint(VAR,VARopt,2);
```

which produces the following output in the Matlab command window:

```
Reduced form VAR estimation:

Real GDP Growth       1-year Int. Rate
constant                        0.3630          -0.0729
Real GDP Growth(-1)             0.3788           0.2607
1-year Int. Rate(-1)            0.0041           0.9541

VAR eigenvalues:
0.3769
0.95592

Reduced-form covariance matrix:
0.28909       0.078151
0.078151      0.14726
```

The cell array `beta` includes the estimated values of the coefficients in the VAR, while the cell array `TABLE` includes – in addition to the estimated coefficients – their standard errors and associated t-statistics.

5 From Reduced-form VARs to Structural VARs

The reduced-form VAR(1) defined by (2) and estimated in the previous section can be written in its structural form as:

$$x_t = c + \Phi x_{t-1} + B\varepsilon_t \tag{6}$$

where x_t, c, Φ have been already defined above; B is a $k \times k$ matrix of coefficients, typically referred to as **structural impact matrix**; and ε_t is a $2 \times T$ matrix of serially uncorrelated innovations, generally referred to as **structural shocks**, which are assumed to be mutually uncorrelated with zero mean and unit variance.[6] The relation

[6]Note that the fact that the variance of the structural shocks is equal to one is just a harmless normalization which does not involve a loss of generality (as long as the diagonal elements of B remain unrestricted. An al-

between the structural shocks and the reduced form innovations is therefore given by the following identity:

$$u_t = B\varepsilon_t, \tag{7}$$

The system of equations defined by the structural VAR (6) should be thought as approximating the true (and unobserved) structure of the economy (for example, the structure of a DSGE model); and the structural shocks as having a well-defined economic interpretation (for example, TFP shocks or monetary policy shocks).

In our simple example of a bivariate VAR(1), let the only the two structural shocks be a demand shock $\left(\varepsilon_t^{Demand}\right)$ and a monetary policy shock $\left(\varepsilon_t^{MonPol}\right)$. This is, of course, another unrealistic simplifying assumption. In reality, there are many more shocks that drive movements in GDP growth and the short-term interest rate. But, as it will become clear below, this assumption is going to simplify the exposition of the identification problem described below, and to clarify the difference between various identification approaches.

The simple structural VAR(1) can be written as a system of linear equations:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix} \tag{8}$$

or:

$$\begin{aligned} y_t &= \phi_{11} y_{t-1} + \phi_{12} r_{t-1} + b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol} \\ r_t &= \phi_{21} y_{t-1} + \phi_{22} r_{t-1} + b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol} \end{aligned} \tag{9}$$

Moreover, as $\varepsilon_t = \left(\varepsilon_t^{Demand\prime}, \varepsilon_t^{MonPol\prime}\right)'$ is assumed to be a $2 \times T$ matrix of uncorrelated white noise processes, their covariance matrix can be written as:

$$\mathbb{V}(\varepsilon_t) \equiv \Sigma_\varepsilon = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2. \tag{10}$$

ternative (and equivalently valid) normalization would be to leave unrestricted the variance of the structural innovations, namely $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2)$ and assume that the diagonal elements of B are equal to 1.

The assumption that the elements of ε_t are mutually uncorrelated is crucial. It implies that we can track the effects that a shock to, say, ε_t^{Demand} has on all variables in the VAR keeping the other shock to zero (and vice versa). The B matrix is also crucial. To see that, consider a unit surprise in ε_t^{MonPol}, i.e. a surprise tightening in monetary policy. What are the consequences for output growth y_t and the short-term interest rate r_t? The answer to this question is given by the second column of the B matrix: y_t will move by b_{12} and r_t will move by b_{22}. This is why the B matrix is also known as the structural impact matrix. The Φ matrix can then be used to track the dynamic effects of the shocks in $t+1$, $t+2$, etc.

The structural innovations ε_t are unobservable, which means that we cannot directly estimate (9). However, we can link the structural innovations and impact matrix to the reduced-form innovations using (7):

$$
\begin{aligned}
u_{yt} &= b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{MonPol}, \\
u_{rt} &= b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{MonPol}.
\end{aligned}
\tag{11}
$$

Equation (11) shows how the reduced-form innovations $u_t = \left(u'_{yt}, u'_{rt}\right)'$ are a linear combination of the structural innovations. Thus, differently from structural VARs, reduced-form VAR cannot be informative about how shocks propagate through the system. An innovation to u_{yt} could be driven by either ε_t^{Demand} or ε_t^{MonPol} (and vice versa). To be able to talk about the causal effects of a shock to the variables in the VAR we need to find a way to recover the B matrix. This is the essence of identification in VARs, which is discussed next.

6 The Identification Problem

The key difference between the structural and reduced-form VARs lies in the covariance matrix of their innovations. While the covariance matrix of the structural VAR innovations is diagonal ($\Sigma_\varepsilon = I_2$), in general the reduced-form innovations are correlated among themselves, so that their covariance is given by a symmetric matrix non-diagonal matrix (Σ_u).

As hinted above, the covariance of the estimated reduced-form residuals plays an important role in VARs because it collects the information on the contemporaneous relation among the variables in the structural system, which (as we have just seen) is also captured by the B matrix. Indeed, using (7) the covariance matrix of the reduced for residuals can be re-written as:

$$\Sigma_u = \mathbb{E}\left[u_t u_t'\right] = B\mathbb{E}\left[\varepsilon_t \varepsilon_t'\right] B' = BB' = \begin{bmatrix} b_{11}^2 & b_{11}b_{21} + b_{12}b_{22} \\ b_{11}b_{21} + b_{12}b_{22} & b_{22}^2 \end{bmatrix} \tag{12}$$

This means that there is a mapping between the estimated covariance matrix of the reduced-form residuals (Σ_u) and the unobserved matrix of structural impact coefficients (B). The identification problem simply boils down to finding a B matrix that satisfies $\Sigma_u = BB'$.

Unfortunately this is not as easy as it sounds. We can think of (**??**) as a system of non-linear equations in the 4 unknown coefficients of the B matrix. The problem is that the Σ_u matrix, given its symmetric nature, leads to only 3 independent restrictions. In other words, we have:

$$\begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ \sigma_{yr}^2 & \sigma_r^2 \end{bmatrix} = \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}}_{B'}, \tag{13}$$

which can be rewritten as the following system of equations:

$$\begin{cases} \sigma_y^2 = b_{11}^2 + b_{12}^2 \\ \sigma_{yr}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_{yr}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_r^2 = b_{21}^2 + b_{22}^2 \end{cases} \tag{14}$$

Because of the symmetry of the Σ_u matrix, the second and the third equation are identical. This means that we are left with 4 unknowns (the b's) but only 3 equations. The system is under-identified, meaning that there are infinite combination of the b's that solve the system of equations (14).

How to solve a system of 3 equations in 4 unknowns? The solution is (typically) to draw from economic theory an additional condition that allows us to recover a fourth equation – and therefore, solve the system of equations (14). There are many ways of solving the identification problem described above. The next section shows how to implement a few popular identification schemes in the VAR Toolbox.

7 Identification in the VAR Toolbox

Many solutions have been developed in the literature to address the identification problem described in the previous section. This section describes how to implement some of the most popular ones by means of simple examples using the VAR Toolbox. Specifically, the next section covers identification by zero contemporaneous restrictions, identification by zero long-run restrictions, identification by sign restrictions, identification with external instruments, and identification with a combination of sign restrictions and external instruments. Appendix A provides the technical details of each identification scheme.

7.1 Identification by zero contemporaneous restrictions

Identification using zero contemporaneous restrictions – also improperly known as Cholesky or recursive identification – was developed by Sims (1980), and is by far the most commonly used identification scheme used in the literature. The idea behind this approach to identification is that some structural shocks may take time to transmit through the economy, and therefore have no contemporaneous effects on (some of) the endogenous variables in the VAR. For example, it is widely believed that there are substantial lags in the transmission of monetary policy to the real economy, while this is not the case for other shocks (such as technology shocks, for example). Under this assumption, one could then impose that monetary policy shocks have zero contemporaneous effects on one (or a subset) of the endogenous variables in the VAR.

This intuition can be formalized in the context of the simple bivariate VAR described in the previous section:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}. \tag{15}$$

As discussed above, the VAR is not identified, as we have four unknowns (the elements of the B matrix) but only three independent equations. Imposing zero contemporaneous restrictions amounts to assuming that some of the non-diagonal elements of B are equal to zero, thus reducing the number of unknowns coefficients.

In the simple case considered here, it will therefore suffice to set to zero one element of the B matrix to be able to solve for the remaining three elements. But which element of the B matrix should be set to zero? One could maintain the assumption that monetary policy shock affects on impact the short-term interest rate (i.e. r_t) but take time to transmit to the real economy and affect output (y_t). This identifying assumption implies that $b_{12} = 0$, so that the structural VAR can be written as:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}, \tag{16}$$

where note that y_t is not contemporaneously affected by ε_t^{MonPol}, while r_t is contemporaneously affected by both ε_t^{Demand} and ε_t^{MonPol}, through the coefficient b_{21} and b_{22}.[7] As we now have 3 independent equations and 3 unknowns, we can recover the elements of the B matrix by solving the system of equations implied by

[7] Typically, it is assumed that the first variable in the system is the only one that is contemporaneously affected by the first structural shock, the second variable is contemporaneously affected by the first and second structural shocks, and so on. The reason for this assumption depends on how this type of identification is commonly implemented, i.e. via a Cholesky decomposition of the reduced-from covariance matrix Σ_u. This assumption is also maintained in the VAR Toolbox. See Appendix A for details.

$\Sigma_u = BB'$. The structural VAR is thus identified.

Zero short-run restrictions can be implemented in the VAR Toolbox with a few lines of code. The structure `VARopt` includes a field that allows the user to choose what identification scheme to employ. The mnemonic for the identification by short-run restrictions is the string `'short'`. So, zero contemporaneous (or short-run) restrictions can be selected by simply executing the following line of code:

```
% Update the VARopt structure to select zero short-run restrictions
VARopt.ident = 'short';
```

Note that the field `VARopt.ident` is automatically set to `'short'` when it is first created as an output of the `VARmodel.m` function – so that, in the case of zero short-run restrictions, the above line of code is actually redundant. It is also useful (but not necessary) to update the `VARopt` structure with a few addition details that will be used when plotting the impulse responses and saving them:

```
% Update the VARopt structure with additional details
VARopt.vnames = Xvnames_long;                % variable names in plots
VARopt.nsteps = 12;                          % max horizon of IRF
VARopt.FigSize = [26,12];                    % size of window for figure
VARopt.firstdate = datesnum(1);              % first date in plots
VARopt.frequency = 'q';                      % frequency of the data
VARopt.snames = {'\epsilon^{Demand}',...     % shock names
    '\epsilon^{MonPol}'};
```

Finally, the actual calculation of the elements of the B matrix under zero short-run restrictions is implemented with the function `VARir.m`, which also computes the impulse responses to the identified shocks.[8] This function takes as an input the estimated `VAR` structure (as described in the previous section) as well as the updated `VARopt` structure, namely:

[8] In a similar fashion, identification can be achieved with the functions for the calculations of forecast error variance decompositions `VARvd.m` and historical decompositions `VARhd.m`.

```
% Compute impulse response
[IR, VAR] = VARir(VAR,VARopt);
```

The `VARir.m` function has two outputs. The first one (`IR`, which we are going to ignore for the moment as it will be the focus of the following sections) is a matrix with the impulse responses calculated according to the identification scheme chosen. The second output is, again, the structure `VAR`, which also served as an input to the function. This is because the `VARir.m` function updates the `VAR.B` field from an empty matrix to the B matrix corresponding to the chosen identification scheme. The B matrix can be printed at screen by executing the following command in the Matlab command window:

```
>> disp(VAR.B)
        0.5377         0
        0.1454    0.3552
```

As discussed above in section 5, the B matrix is crucial to be able to track the effects of a shock through the system. Consider a unit surprise in ε_t^{MonPol}, i.e. a surprise tightening in monetary policy, in the structural VAR in equation (16). What are the impact effects of such shock on output growth y_t and the short-term safe rate r_t? The answer to this question is given by the second column of the B matrix. To see that:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix} = \begin{bmatrix} 0.5377 & 0 \\ 0.1454 & 0.3552 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.3552 \end{bmatrix}$$

The response of output is equal to 0 – not surprisingly, as we assumed so. The response of the interest rate is instead equal to 0.3552. These impact effects then propagate through the system over time according to the transition matrix Φ. For example, the effect of the shock on the endogenous variables one quarter after the shock has hit is given by:

$$\begin{bmatrix} y_{t+1} \\ r_{t+1} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \texttt{0.3788} & \texttt{0.0041} \\ \texttt{0.2607} & \texttt{0.9541} \end{bmatrix} \begin{bmatrix} \texttt{0} \\ \texttt{0.3552} \end{bmatrix} = \begin{bmatrix} \texttt{0.0015} \\ \texttt{0.3388} \end{bmatrix}$$

To see that, execute the following command in the Matlab command window:

```
>> disp(VAR.Fcomp*VAR.B(:,2))
        0.0015
        0.3388
```

This is called the impulse response function at horizon $h = 2$. The effect of the shock in $h = 3$, $h = 4$, etc. can be computed in a similar fashion, i.e. using again the transition matrix Φ to iterate forward the values of the endogenous variables from horizon $h = 2$ to horizon $h = 3$, $h = 3$ to $h = 4$, etc.[9] The resulting impulse response function is stored in the matrix `IR`. In this example, the `IR` matrix has dimension 12 (as specified above in `VARopt.nsteps`) $\times 2$ (the number of endogenous variables, y_t and r_t) $\times 2$ (the number of shocks, ε_t^{Demand} and ε_t^{MonPol}). So, for example, the response of output growth and the interest rate to the monetary policy shock in the first year since the shock hit can be printed at screen by executing the following command:

```
>> disp(IR(1:4,:,2))
        0          0.3552
        0.0015     0.3388
        0.0019     0.3237
        0.0021     0.3093
```

[9] A detailed description of the calculation of impulse responses is provided in Appendix B.

NOTE As explained in detail in Appendix A, this identification is achieved with a Cholesky decomposition of the reduced form covariance matrix. This implies that the ordering of the variables in the matrix `X` matters. The `'short'` option implicitly assumes that the `B` matrix is lower triangular. In turn this means that the structural shock associated with the first equation affects all variables in the system (as captured by the first column of the `B` matrix); the the structural shocks associated with the second equation has zero effect on the first endogenous variable, and affects the variables in the system; etc.

Finally, note that a time series of the structural shocks can be obtained by inverting equation (7), which gives $\varepsilon_t = B^{-1}u_t$. In Matlab, the structural shocks can be computed by executing a single line of code:

```
% Compute structural shocks (Tx2)
eps_short = (VAR.B\VAR.resid')';
```

where `eps_short` is the $T \times 2$ matrix of structural shocks. It is also possible to verify that the structural shocks are orthogonal to each other by typing in the command window:

```
>> disp(corr(eps'))
        1.0000    0.0000
        0.0000    1.0000
```

which implies that the covariance matrix of the structural shocks is diagonal.

7.2 Identification by zero long-run restrictions

Blanchard and Quah (1989) proposed an alternative identification method that builds on a similar intuition to the zero short-run restrictions described in the previous section, but imposes zero restrictions on the long-run effect of structural shocks. For example, some models imply that only technology shocks have a permanent effect on the level of output. On the contrary, demand shocks have a zero effect on the level of output in the long-run. Under this assumption, one could then impose that non-technology shocks have zero contemporaneous effects on the level of output in the long-run.

But how to map the B matrix, which captures the contemporaneous effects of structural shocks on the endogenous variables in the VAR, to the long-run effects of structural shocks? To give intuition in the context of the simple bivariate VAR described in the previous section, assume that the two shocks driving the model economy are a demand shock (ε_t^{Demand}) and a technology shock (ε_t^{Supply}):

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} c_y \\ c_r \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Supply} \\ \varepsilon_t^{Demand} \end{bmatrix}, \tag{17}$$

The effect of these structural shocks on the endogenous variables at different horizons h can be computed easily, as we have seen in the previous section, by recursively iterating forward the structural (and yet unobserved) impact matrix, B. That is:

$$\begin{aligned} & x_t = B\varepsilon_t, \\ & x_{t+1} = \Phi B\varepsilon_t, \\ & ... \\ & x_{t+\infty} = \Phi^{\infty} B\varepsilon_t. \end{aligned} \tag{18}$$

The long-run (i.e. for h that goes to infinity) cumulative effect of the shock can be obtained by summing all the terms in (18):

$$x_{t,t+\infty} = B\varepsilon_t + \Phi B\varepsilon_t + \Phi^2 B\varepsilon_t + ... + \Phi^{\infty} B\varepsilon_t = \sum_{j=0}^{\infty} \Phi^j B\varepsilon_t, \tag{19}$$

where $x_{t,t+\infty}$ denotes the sum from t to $t+\infty$ of the elements in (18). Finally note that, if the VAR is stable (i.e. if the eigenvalues of Φ lie inside the unit circle), the infinite sum in equation (19) converges to:

$$x_{t,t+\infty} = (I - \Phi)^{-1} B\varepsilon_t = C\varepsilon_t, \tag{20}$$

where

$$C \equiv (I - \Phi)^{-1} B \tag{21}$$

is a 2×2 matrix which captures the cumulative effect of shocks ε_t on x_t from time t to $t + \infty$.

The idea behind identification through zero long-run restrictions is to impose a zero restriction on the long-run impact matrix C. For example, one could maintain the assumption that demand shocks have zero effect on the level of output in the long run. Equation (20) allows to exactly impose this restriction by setting $c_{12} = 0$:

$$\begin{bmatrix} y_{t,t+\infty} \\ r_{t,t+\infty} \end{bmatrix} = \begin{bmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Supply} \\ \varepsilon_t^{Demand} \end{bmatrix}. \tag{22}$$

In fact, the upper right element of C captures the long-run *cumulative* effect of ε_t^{Demand} on the growth rate of GDP, i.e. its effect on the level of GDP.

But how does this help with the identification of B? Clearly, $C \equiv (I - \Phi)^{-1} B$ is unknown as B is unobserved. But Equation (21) can be exploited to achieve identification. To see that, define $\Omega \equiv CC'$ and note that Ω is known:

$$\Omega \equiv CC' = \left((I - \Phi)^{-1} \right) BB' \left((I - \Phi)^{-1} \right)' = \left((I - \Phi)^{-1} \right) \Sigma_u \left((I - \Phi)^{-1} \right)'. \tag{23}$$

Equation (23) therefore provides a mapping between the known 2×2 matrix Ω and the unobserved C matrix – and thus the B matrix through equation (21) – which could be used to solve for the unknown elements of B.

The matrix Ω, however, is a positive-definite symmetric matrix, which implies that it provides only three independent restrictions for four unknowns:

$$\begin{bmatrix} \omega_y & \omega_{yr} \\ \omega_{yr} & \omega_r \end{bmatrix} = \begin{bmatrix} c_{11}^2 + c_{12}^2 & c_{11}c_{21} + c_{12}c_{22} \\ c_{11}c_{21} + c_{12}c_{22} & c_{21}^2 + c_{22}^2 \end{bmatrix} \tag{24}$$

In a similar way to the the identification by short-run restrictions, assuming $c_{12} =$

0 (i.e. assuming that the long-run *cumulative* effect of ε_t^{Demand} on the growth rate of GDP is equal to zero) allows to solve the system of equations implied by (24). Finally, once C is known, it is possible to recover the structural impact matrix B using (21) and, thus, to identify the VAR.

Zero long-run restrictions in The VAR Toolbox can be implemented with a few lines of code. As for zero short-run restrictions, the first step is to set the `VARopt.ident` field appropriately. The mnemonic for long-run restrictions is the string `'long'`. So, zero long-run restrictions can be selected by simply running the following line of code:

```
% Update the VARopt structure to select zero long-run restrictions
VARopt.ident = 'long';
```

It is also useful (but not necessary) to update the `VARopt` structure with a few addition details that will be used when plotting the impulse responses, saving them, etc. As most setting have been set for the previous example, it will suffice to run the following lines of code:

```
% Update the VARopt structure with additional details
VARopt.snames = {'\epsilon^{Supply}','\epsilon^{Demand}'};
```

As for the case of other identification schemes, the actual implementation of the zero long-run restrictions is via the `VARir.m`, `VARvd.m`, or `VARhd.m` functions. For example:

```
% Compute impulse responses
[IR, VAR] = VARir(VAR,VARopt);
```

As before, the `VARir.m` function generates a matrix of impulse responses (`IR`) and updates the `VAR` structure with a new `VAR.B` field consistent with the assumption of $c_{12} = 0$. The B matrix can be printed at screen by executing the following command in the Matlab command window:

```
>> disp(VAR.B)
        0.5368   -0.0309
```

```
        0.1655    0.3462
```

Note that the B matrix, which was left unrestricted, is not lower triangular anymore (as in the case of zero short-run restrictions), but has non-zero entries in both columns. As before, each column represents the impact impulse response to the supply and demand shocks, respectively.

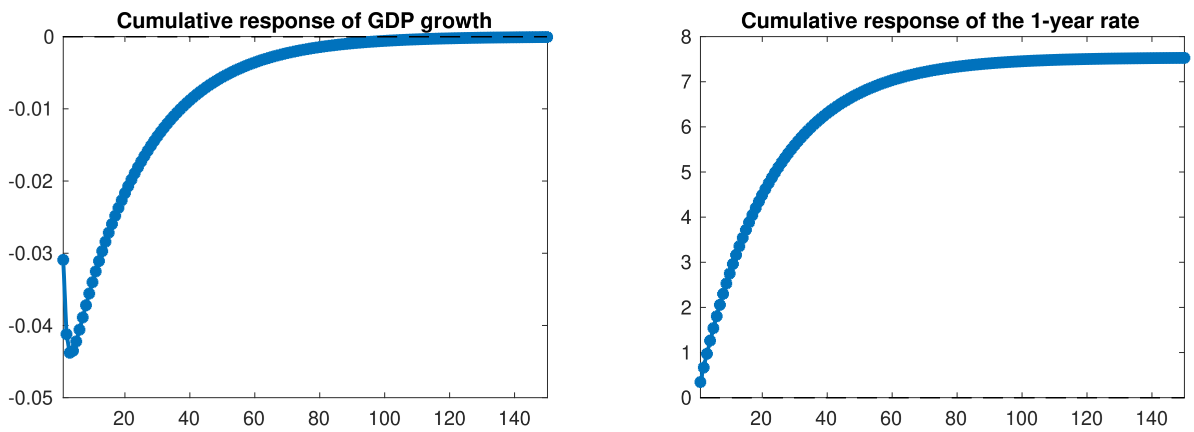
To check that the structural VAR estimated with the above commands is consistent with the assumptions made, it is necessary to compute the C matrix – the matrix that captures the cumulative long-run effect of shocks on the endogenous variables. Under the assumptions made above, the C matrix should be lower triangular, so that the cumulative effect of ε_t^{Demand} on the growth rate of GDP is zero. Recalling from equation (21) that $C \equiv (I - \Phi)^{-1} B$, the C matrix can be printed at screen with command in the Matlab command window:

```
>> disp((eye(2)-VAR.Fcomp)\VAR.B)
        0.9224    0.0000
        8.8389    7.5367
```

As assumed, the c_{12} element is equal to zero. It is also possible to check that the assumed restriction holds by plotting the cumulative response of the endogenous variables to the demand shock at a long enough horizon, and checking that the cumulative impulse response of output growth, y_t, converges to zero. Figure 2 plots the cumulative impulse response functions of the interest rate and output growth to the demand shocks for 150 quarters:

The non-zero impact responses are consistent with the B matrix reported above. In the long-run, the effect of the demand shock on the GDP *level* (i.e. the cumulated GDP growth rates) tends to zero (left panel), while this is not the case for the interest rate (right panel).

Figure 2 CUMULATIVE IMPULSE RESPONSES TO ε_t^{Demand}



NOTE. Cumulative impulse responses of real GDP growth and the 1-year Treasury bill yield to a demand shock identified with zero long-run restrictions. Percentage points.

7.3 Identification by sign restrictions

While the zero restrictions discussed in the previous sections can be justified by economic theory, in many applications these restrictions are implausible or hard to justify. The identification by sign restrictions provides an alternative approach that exploits prior beliefs (typically informed by theroetical models) about the sign that certain shocks should have on certain endogenous variables.

The idea is to impose restrictions on a *set* of orthogonalised impulse response functions. So, differently from the identification schemes described above (where there is a unique point estimate of the B matrix), sign restricted VARs are only set identified. In other words, the data are potentially consistent with a wide range of B matrices that are all admissible in that they satisfy the sign restrictions – see **?**, **?**, and Uhlig (2005).

To fix ideas, a demand shock (ε_t^{Demand}) should lead to an increase in output growth (y_t) and to an increase in the short term interest rate (r_t), as monetary policy responds to the shock by tightening its stance to contain the boom. Differently, a monetary policy shock (ε_t^{MonPol}) should lead to a fall in output growth for an unexpected increase in interest rates. That is:

Table 1 Sign restrictions for demand and monetary policy shocks

	Demand (ε_t^{Demand})	Monetary Policy (ε_t^{MonPol})
Output growth (y_t)	+	-
Short-rate Int. Rate(r_t)	+	+

The signs in the above table represent restrictions on the elements of the structural impact matrix B. But how can such restrictions be imposed? The key intuition for the sign restrictions identification is based on the following three steps.

1. Orthonormal matrices (Q). An orthonormal matrix Q is a real square matrix whose columns and rows are orthogonal unit vectors.[10] What does it mean? Take for example two 2×1 vectors q_1 and q_2, then the matrix $Q = (q_1, q_2)$ is orthonormal if (i) the vectors have unit norm ($\| q_i \| = 1$), and (ii) the vectors are mutually orthogonal ($q_1^T q_2 = 0$). It follows that

$$QQ' = I_2 \quad \text{and} \quad Q' = Q^{-1}$$

Note that it is possible to generate a large number of matrices that satisfy the above conditions by computing the orthogonal factor in the QR factorization of a random matrix with elements from the standard normal distribution – see the `getqr.m` function and the examples therein.

2. Candidate structural impact matrices (B_j). Consider the structural impact matrix B corresponding to the Cholesky factor of the reduced form covariance matrix Σ_u of our simple bivariate VAR, namely:

$$\Sigma_u = PP'.$$

We know from Section 7.1 that $B = P$ is the unique structural impact matrix that would obtain under the zero contemporaneous restriction $b_{12} = 0$. Now note that

[10]More precisely, Q is a matrix distributed according to the Haar measure over the group of orthogonal matrices.

the following equality holds

$$\Sigma_u = PP' = PQ_jQ_j'P' = \underbrace{(PQ_j)}_{B_j}\underbrace{(PQ_j)'}_{B_j'} \tag{25}$$

where Q_j denotes a randomly drawn orthonormal matrix such that $Q_jQ_j' = I_2$. The key property of B_j is that, in addition to satisfying (25), it is such that the associated structural shocks $\varepsilon_{jt} = B_j u_t$ are orthogonal and have unit variance. That is, $\Sigma_{j,\varepsilon} = I_2$. It follows that B_j is a valid 'candidate' structural impact matrix that solves the identification problem.

Also note that B_j is not triangular any more. As there are infinite of these B_j matrices, which one should we use? How do we know whether B_j represents a plausible solution?

3. Checking the sign restrictions. The idea of sign restrictions is to generate a large enough number of B_j matrices and retain those that satisfy a set of *a priori* signs for the response of the endogenous variables – where recall that the B_j matrix contains the impact response of all endogenous variables to all structural shocks. For example, for a given Q_j matrix and associated structural impact matrix B_j, the structural representation of our VAR can be written as:

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{j,11} & b_{j,12} \\ b_{j,21} & b_{j,22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}, \tag{26}$$

where the elements of B_j are known and equal to PQ_j. We can then check whether the impact response of the two structural shocks to output growth and the short-term interest rate satisfy the sign restrictions:

	Demand (ε_t^{Demand})	Monetary Policy (ε_t^{MonPol})
Output growth (y_t)	$b_{j,11} > 0$?	$b_{j,12} < 0$?
Short-rate Int. Rate(r_t)	$b_{j,21} > 0$?	$b_{j,22} > 0$?

If all the elements of B_j satisfy the sign restrictions, then we retain the draw and store PQ_j. If at least one element of B_j does not satisfy the restrictions, we discard the draw, compute a new Q_j matrix, and check again the signs of the associated new B_j. After drawing a large number of Q_j matrices that satisfy the sign restrictions, we can then construct a distribution of the B_j matrix – as well as of the impulse responses variance decompositions, etc. Importantly, as noted above, VARs identified with sign restrictions are only *set-identified*. In other words, the data are potentially consistent with a wide range of structural models that are all admissible in that they satisfy the identifying restrictions.

Sign restrictions in the VAR Toolbox can be implemented with a few lines of code. The key element is to specify the sign restrictions. This is done with a square matrix with size equal to the number of endogenous variables – i.e., in the case of the examples in this section a 2×2 matrix – containing 1s, which stand for a positive response; -1s, which stand for a negative response; and 0s which stand for an unrestricted response. Consistent with the notation in equation (26), each row contains restrictions for a given structural shock. Thus, the restrictions described in Table 1 can therefore be implemented in Matlab with the following line of code:

```
% Define sign restrictions : positive 1, negative -1, unrestricted 0
SIGN = [ 1, 1  ; % Real GDP
        -1, 1]; % 1-year rate
% Update the VARopt structure with inputs to the sign restriction
routine
VARopt.ndraws = 500;
VARopt.sr_hor = 1;
VARopt.pctg = 68;
```

The matrix `SIGN` specifies the sign restrictions as in Table 1. It is also possible, though not necessary to update some of the fields in the VARopt structure. The field `ndraws` specifies the number of accepted draws the routine needs to find. The field `sr_hor` specifies the number of periods the sign restrictions specified in are required to hold. In this specific example, we have set the restrictions to hold for one quarter only (namely, on impact), but it is possible to specify restrictions to hold for longer horizons. The field `pctg` specifies the confidence levels for the credible intervals.

As for other identification schemes, it is also useful (but not necessary) to update the `VARopt` structure with a few addition details that will be used when plotting the impulse responses, saving them, etc. As most setting have been set for the previous example, it will suffice to run the following lines of code:

```
% Update the VARopt structure with additional details
VARopt.figname= 'graphics/sign_';          % folder and file prefix
VARopt.FigSize = [26 8];                   % size of window for figure
VARopt.snames = {'\epsilon^{Demand}',...  % shocks names
     '\epsilon^{MonPol}'};
```

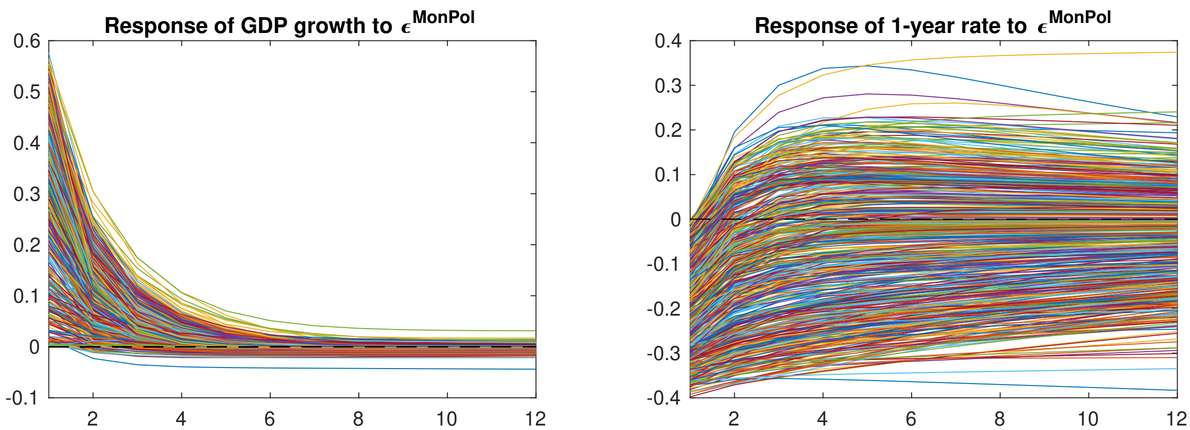
Note that, differently from the other identification schemes, it is not required to change the `VARopt.ident` field to the sign restrictions mnemonic. This is because, unlike the other identification schemes, the sign restrictions procedure is not implemented with the `VARir.m`, `VARvd.m`, or `VARhd.m` functions, but rather with the `SR.m` function as follows:

```
% Implement sign restrictions identification with SR routine
SRout = SR(VAR,SIGN,VARopt);
```

The structure `SRout` contains all relevant output from the sign restriction procedure. Of particular interest for the discussion in this section are two matrices. The matrix `SRout.Ball` includes all the accepted draws of B_j, and thus has a dimension of $2 \times 2 \times 500$. Each of the accepted B_j, which by definition satisfies the sign restrictions in Table 1, is associated with an impulse response function, stored in the matrix `SRout.IRall`. Figure 6 reports all the impulse responses of GDP growth and the 1-year rate to the monetary policy shock.

While all accepted draws are associated with a structural representation of the VAR that satisfies the identifying restrictions, it is common to report a summary measure of the identified set. The matrix `SRout.Bmed` is computed as the median

Figure 3 IMPULSE RESPONSES TO A MONETARY POLICY SHOCK



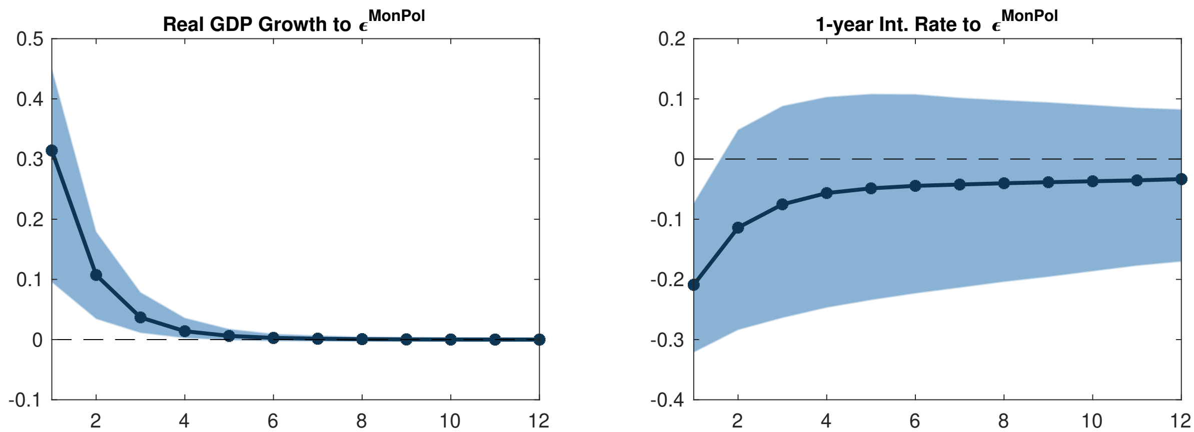
NOTE. Full identified set of impulse responses of real GDP growth and the 1-year Treasury bill yield to a monetary policy shock identified with sign restrictions. Percentage points.

of `SRout.Ball` across all accepted draws, and thus has a dimension of 2×2. Similarly, it is possible to compute different quantiles of the distribution of `SRout.Ball` (and the associated `SRout.IRall`) to plot credible intervals for the impulse responses. The 16th and 84th quantiles of the distribution of `SRout.IRall` is stored in the matrices `SRout.IRinf` and `SRout.IRsup`, respectively. The VAR Toolbox has a built-in function (`VARirplot.m`) to plot the impulse responses from any of the identification schemes described in this section. The impulse responses identified with sign restrictions can be plotted with the following line of code:

```
% Plot credible intervals
VARirplot(SRout.IRmed,VARopt,SRout.IRinf,SRout.IRsup)
```

Figure 4 reports the median impulse response of GDP growth and the 1-year Treasury rate to the monetary policy shock, together with 68 Percentage points. credible intervals, as specified above.

Figure 4 IMPULSE RESPONSES TO A MONETARY POLICY SHOCK



NOTE. Impulse responses of real GDP growth and the 1-year Treasury bill yield to a monetary policy shock identified with sign restrictions. Percentage points.

7.4 Identification with external instruments (or proxies)

The external instruments identification approach has been proposed by Stock and Watson (2012) and Mertens and Ravn (2013). This identification strategy uses standard instrumental variable techniques to isolate the variation of the VAR reduced-form residuals that are due to the structural shock of interest. The key element of this identification technique is, thus, the presence of an instrument that is correlated with a structural shock of interest and uncorrelated with all other structural shocks.

For the example in this section, assume that the data are driven by a demand shock, as well as another shock (or a combination of shocks) that we leave unidentified. Also assume that a valid instrument (z_t) for the demand shock exists, namely that z_t is correlated with ε_t^{Demand} and uncorrelated with the other shock ε_t^{Other}. More formally, z_t satisfies the following properties:

$$\mathbb{E}\left[\varepsilon_t^{Demand} z_t'\right] = c, \tag{27}$$

$$\mathbb{E}[\varepsilon_t^{Other} z_t'] = 0, \tag{28}$$

If such an instrument exists, it is possible to identify the contemporaneous response of all endogenous variables to the demand shock. That is, we can identify one column (in this example, the first one) of the B matrix:

$$B = \begin{bmatrix} b_{11} & - \\ b_{21} & - \end{bmatrix} \tag{29}$$

The intuition is as follows. Recall that the reduced-form residuals u_{yt} and u_{rt} are a linear combination of two orthogonal shocks ε_t^{Demand} and ε_t^{Other}:

$$\begin{aligned} u_{yt} &= b_{11}\varepsilon_t^{Demand} + b_{12}\varepsilon_t^{Other}, \\ u_{rt} &= b_{21}\varepsilon_t^{Demand} + b_{22}\varepsilon_t^{Other}. \end{aligned} \tag{30}$$

It is therefore possible to isolate the variation in one of the two reduced-form residuals (say, u_{yt}) that is due only to the shock of interest with a regression of u_{yt} itself on the instrument z_t:

$$u_{yt} = \beta z_t + \xi_t, \tag{31}$$

The fitted values of this first stage regression $\hat{u}_{yt} = \hat{\beta} z_t$ capture the variation of u_{yt} that is due to the demand shock. As z_t is orthogonal to ε_t^{Other}, the variation of u_{yt} that is due to the other shock (namely, the component $b_{12}\varepsilon_t^{Other}$) ends up in the residual ξ_t.

By projecting the residuals of the interest rate equation u_{rt} on the fitted values of the previous regression $\hat{u}_{yt}$, it is possible to get a consistent estimate of the ratio b_{21}/b_{11}

$$u_{rt} = \underbrace{\gamma}_{b_{21}/b_{11}} \hat{u}_{yt} + \zeta_t, \tag{32}$$

Under the assumption that $\mathbb{E}[\varepsilon_t^{Demand} z_t'] = 0$, the fitted values $\hat{u}_{yt}$ are orthogonal to ε_t^{Other}. Therefore, with a similar logic to the first stage regression, the variation in u_{rt} that is due to ε_t^{Other} ends up in the residual ζ_t, while the fitted values $\hat{\gamma}\hat{u}_{yt}$ isolate the variation in u_{rt} that is due to the demand shock.

If we consider the effect of a demand shock that increases GDP growth by 1 percentage point (i.e. by normalizing $b_{11} = 1$) we can easily recover $b_{21} = \gamma$.[11] The procedure described in this section thus provides an estimate of the first column of B up to a scaling factor:

$$B = \begin{bmatrix} 1 & - \\ \gamma & - \end{bmatrix} \tag{33}$$

which can then be used to compute impulse response to the demand shock as described above.

The first step to implement the identification with external instruments in VAR Toolbox is to add the time series of the instrument to the `VAR` structure. Typically, the instrument exploits 'external' information that is independent of the VAR, such as narrative fiscal shocks as in Mertens and Ravn (2013) or high frequency monetary policy surprises as in Gertler and Karadi (2015). For simplicity, the example in this section exploits an instrument for the demand shock (stored in the column vector `iv`) that is artificially constructed by adding some noise to the demand shock identified in the zero short-run restrictions example. The code below shows how to update the `VAR` structure:

```
% Update VAR structure with external instrument
VAR.IV = iv;
```

As for the short- and long-run zero restrictions identification schemes, the second step is to set the `VARopt.ident` structure appropriately. The mnemonic for the external instruments identification is the string `'iv'`. It is also useful (but not necessary) to update the `VARopt` structure with a few additional details that will be used when plotting the impulse responses and saving them:

```
% Update the options in VARopt
VARopt.ident = 'iv';
VARopt.snames = {'\epsilon^{Demand}','\epsilon^{Other}'};
```

The actual implementation of the external instruments identification restrictions is

[11] In the VAR toolbox, the impulse response are further normalized to match the standard deviations of the shocks, as explained in Gertler and Karadi (2015). See Appendix 7 for details.

via the `VARir.m` function:

```
% Compute impulse responses
[IR, VAR] = VARir(VAR,VARopt);
```

As before, the `VARir.m` function updates the `VAR` structure with a new `VAR.B` field consistent with the chosen identification scheme. Differently from the previous examples though, it also updates the `VAR` structure with an additional structure including some information about the first stage regression (`VAR.FirstStage`). As in a standard instrumental variable approach, the F-statistic of the first stage regression is important to assess the relevance of the instrument.

As discussed above, only the first column of the B matrix is identified, while the second column is set to zero by construction. The B matrix can be printed at screen by executing the following command in the Matlab command window:

```
>> disp(VAR.B)
        0.5375    0
        0.1538    0
```

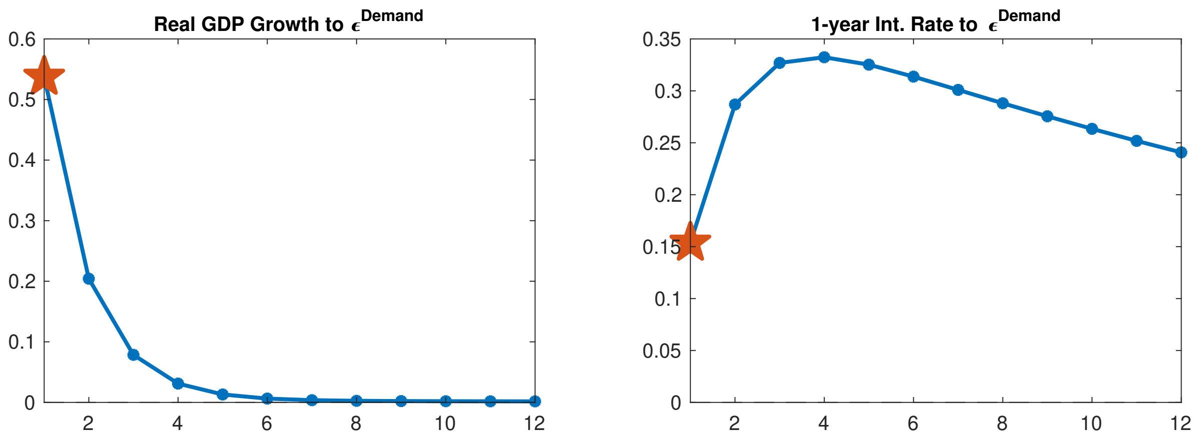
Once the impact responses are obtained with the first stage and second stage regressions (reported in Figure 5 with stars), the impulse responses at horizons $h > 1$ to the demand shock are then computed as usual and stored in the matrix (`IR`). Figure 5 reports the impulse response to the identified demand shock

7.5 Identification by a combination of sign restrictions and external instruments

The external instruments and sign restriction identification approaches can be combined as proposed by Cesa-Bianchi and Sokol (2021). The main idea of this approach is to employ external instruments to identify one (or more) shocks and the remaining shocks (or a subset of them) with sign restrictions.

This section builds on the example used in the previous section, where a demand shock is identified with external instruments and a second shock is left un-

Figure 5 Impulse Responses to a Demand Shock



Note. Impulse responses of real GDP growth and the 1-year Treasury bill yield to a demand shock identified with external instruments. Percentage points.

identified. Differently from the previous section, however, assume here that the second shock driving the data is a monetary policy shock. As the demand shock can be identified with external instruments, we only need to specify the sign restrictions for the monetary policy shock as discussed in section 7.3. Table 2 summarizes the identifying restrictions on the elements of the structural impact matrix B.

Table 2 Restrictions for demand and monetary policy shocks

	Demand (ε_t^{Demand})	Monetary Policy (ε_t^{MonPol})
Output growth (y_t)	Ext. Instrument	-
Short-rate Int. Rate(r_t)	Ext. Instrument	+

To see how to combine the external instrument and sign restrictions approaches, start by partitioning the matrix B into a column vector b, which captures the impact of the demand shock, and a column vector B^{SR}, which captures the impact of the monetary policy shock:

$$B = \left[\begin{array}{cc} B^{IV} & B^{SR} \end{array} \right]. \tag{34}$$

where B^{IV} and B^{SR} are 2×1 vectors.[12] Assuming that a valid instrument for the

[12]Note that both B^{IV} and B^{SR} can be matrices, i.e. can include more than one shock, as in the original paper by

demand shock exists, the first column of the B matrix (B^{IV}) can be easily identified as explained in the previous section.

We now show how to combine the external instruments identification approach with a standard sign restriction approach to identify the remaining structural shock (ε_t^{MonPol}) conditional on the shock identified with the external instrument (ε_t^{Demand}). To identify B^{SR} (i.e. the contemporaneous impact of the monetary policy shocks) we proceed as follows. First, using (34), re-write the covariance matrix of the reduced-form residuals as:

$$\Sigma_u = BB' = \left[\begin{array}{cc} B^{IV} & B^{SR} \end{array} \right] \left[\begin{array}{cc} B^{IV} & B^{SR} \end{array} \right]'. \tag{35}$$

As we seen above, this decomposition of the covariance matrix is not unique. Let P be the Cholesky decomposition of the covariance matrix Σ_u, and let Q_j be a randomly drawn orthonormal matrix (where, as before, j denotes a random draw) such that $Q_j Q_j' = I_2$, then:

$$\Sigma_u = PP' = PQ_jQ_j'P' = \left(PQ_j\right)\left(PQ_j\right)' \tag{36}$$

Similarly to the sign restriction procedure, the identification strategy described in this section consists in constructing a large number of orthonormal matrices Q_j that satisfy the following condition:

$$PQ_j = \left[\begin{array}{cc} B^{IV} & B_j^{SR} \end{array} \right] \tag{37}$$

where B^{IV} is point identified with the external instrument and B_j^{SR} is set identified with the sign restrictions. Thus, the main difference with the standard sign restriction procedure lies in the construction of the Q_j matrix. Instead of obtaining Q_j from a QR factorization of a random matrix with elements from the standard normal distribution, here we construct the Q_j matrix sequentially, with the following two steps:

Cesa-Bianchi and Sokol (2021).

1. Find a normal vector Q^{IV} of dimension 2×1 that rotates the first column of P (the Cholesky decomposition of Σ_u) into the vector B^{IV}. That is, we find a $n \times 1$ normal vector Q^{IV} such that the following equality holds:

$$PQ^{IV} = B^{IV} \tag{38}$$

2. Given Q^{IV}, build the remaining column of an orthonormal 2×2 matrix Q_j following a standard Gram-Schmidt process.[13] That is, find a (2×1) vector Q_j^{SR} such that the following equality holds:

$$\begin{bmatrix} Q^{IV} & Q_j^{SR} \end{bmatrix} \begin{bmatrix} Q^{IV} & Q_j^{SR} \end{bmatrix}' = Q_j Q_j' = I. \tag{39}$$

As in the standard sign restriction procedure, the matrix $B_j = PQ_j$ then represents a candidate structural representation because:

$$\Sigma_u = \begin{bmatrix} B^{IV} & B_j^{SR} \end{bmatrix} \begin{bmatrix} B^{IV} & B_j^{SR} \end{bmatrix}' = \underbrace{P \begin{bmatrix} Q^{IV} & Q_j^{SR} \end{bmatrix}}_{B_j} \underbrace{\begin{bmatrix} Q^{IV} & Q_j^{SR} \end{bmatrix}' P'}_{B_j'} \tag{40}$$

and because $B_j = PQ_j$ is such that the associated structural shocks $\varepsilon_t = Bu_t$ are orthogonal and have unit variance. It is therefore possible to check whether the elements of B_j associated with a given random matrix Q_j are consistent with the restrictions in Table 2 – and, if so, retain the draw.

The first step for the implementation of the identification with external instruments and sign restrictions in VAR Toolbox is to identify the first column of the B matrix with the external instruments approach, as discussed in the previous section. Note that, in addition to the various outputs described above, the `VARir.m` function stores the B^{IV} vector (i.e. the first column of the B matrix) in `VAR.Biv`. The B^{IV} vector can be printed at screen by executing the following command in the Matlab command window:

[13] Additional details on how to construct the remaining columns of Q are reported in Appendix A.

```
>> disp(VAR.Biv)
        0.5375
        0.1538
```

The following step consist in defining the sign restrictions. Differently from the standard sign restrictions approach, here it is only necessary to specify the sign restrictions for the shocks that are not identified with external instruments – in the case of the simple bivariate VAR considered in this section, the monetary policy shock. As before, this is done by specifying a matrix `SIGN` containing the restrictions. The size of the `SIGN` matrix has to be equal to the size of the B^{SR} matrix, i.e. the number of rows has to be equal to the number of endogenous variables and the number of columns equal to the number of shocks identified with sign restrictions. In the case of the example in this section, `SIGN` is therefore a 2×1 matrix:

```
% Define sign restrictions to identify monetary policy shock
% Positive 1, Negative -1, Unrestricted 0:
SIGN = [ 1; % Real GDP
        -1]; % 1-year rate
```

The matrix `SIGN` specifies sign restrictions consistent with those in the second column of Table 2. As for other identification schemes, it is also useful (but not necessary) to update the `VARopt` structure with a few addition details that will be used when plotting the impulse responses, saving them, etc. As most setting have been set for the previous example, it will suffice to run the following lines of code:

```
% Update the VARopt structure with additional details
VARopt.figname= 'graphics/iv_sign_';
VARopt.snames = {'\epsilon^{Demand}','\epsilon^{MonPol}'};
```

Finally, the identification with sign restrictions can be implemented as in the previous example, with the `SR.m` function as follows:

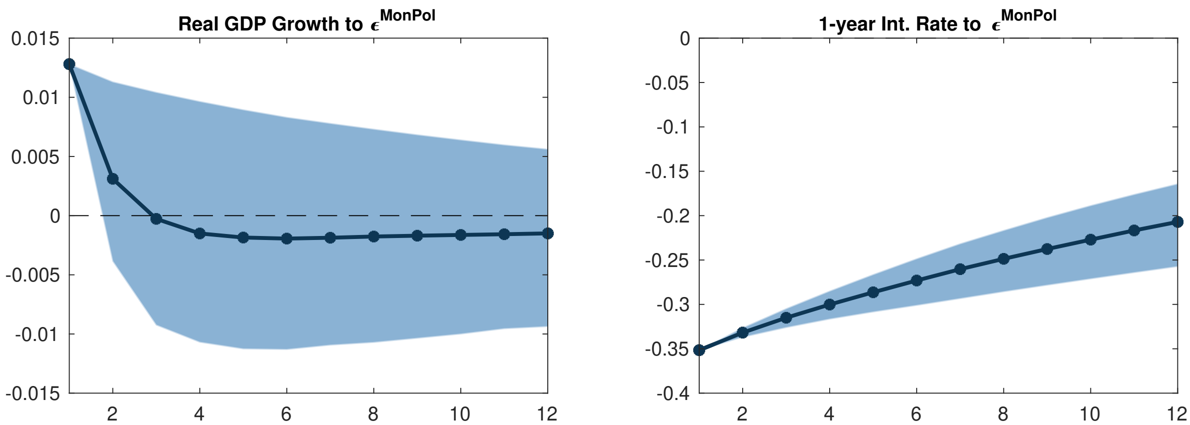
```
% Implement sign restrictions identification with SR routine  %
conditional on VAR.Biv being already identified
SRIVout = SR(VAR,SIGN,VARopt);
```

The structure `SRIVout` contains all relevant output from the mix of external instruments and sign restriction procedure. As in teh sign restriction example of section XX, the matrix `SRout.Ball` includes all the accepted draws of B_j, and thus has a dimension of $2 \times 2 \times 500$. Importantly though, and differently from the stadard sign restrictions procedure, the first column of B_j is always equal to B^{IV}, namely the estimates from the external instruments approach. One way to check this is by printing at screen the median of the B_j matrices across the $j = 500$ draws (stored in `SRIVout.Bmed`), which should be equal to the vector `VAR.Biv`, as shown with the following line of code:

```
>> disp(SRIVout.Bmed)
        0.5375    0.0128
        0.1538   -0.3516
```

Finally, as for the standard sign restriction approach, the impulse responses (and associated credible bands) are stored in the matrices `SRIVout.IRmed`, `SRIVout.IRinf`, and `SRIVout.IRupp`. As the impulse responses to the demand shock are, by construction, equal to those in the external instrument example of section 7.4, Figure 6 reports the impulse responses of GDP growth and the 1-year rate to the monetary policy shock.

Figure 6 IMPULSE RESPONSES TO A MONETARY POLICY SHOCK



NOTE. Full identified set of impulse responses of real GDP growth and the 1-year Treasury bill yield to a monetary policy shock identified with sign restrictions. Percentage points.

8 Conclusions

The main objective of this paper is to provide intuition for the mechanics of VAR models by means of a series of practical examples implemented with the VAR Toolbox. The VAR Toolbox is a collection of Matlab routines that, in a consistent way, allows to perform standard VAR analysis, such as the estimation of VAR models, the identification of structural shocks, and the computation of impulse responses, forecast error variance decompositions, and historical decompositions.

The VAR Toolbox, and this accompanying handbook, is targeted at users who are not familiar with VAR models and want to get an informal intuition behind their workings. Most of the examples presented in this handbook are overly simplistic from an economic standpoint, but have the advantage of being easy to follow and to bring to the computer. Therefore, this handbook is not a substitute to standard time series econometrics textbooks, but rather a complement – which is hopefully making easier for users to follow more comprehensive and formal treatments of these topics.

The VAR Toolbox should also be (hopefully) useful to applied researchers who want to perform standard VAR analysis or to extend common approaches for research purposes. All the codes are public, and users are indeed encouraged to modify them as they wish for their own research purposes. If you do that, please get in touch or raise an issue in Github so that the VAR Toolbox can keep on evolving.

References

BLANCHARD, O. J. AND D. QUAH (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–73.

CANOVA, F. AND F. FERRONI (2020): "A hitchhiker guide to empirical macro models," CEPR Discussion Papers 15446, C.E.P.R. Discussion Papers.

CESA-BIANCHI, A. AND A. SOKOL (2021): "Financial shocks, credit spreads, and the international credit channel," *Journal of International Economics*, 1035–43.

GERTLER, M. AND P. KARADI (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

MERTENS, K. AND M. O. RAVN (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, 103, 1212–47.

RUBIO-RAMIREZ, J. F., D. F. WAGGONER, AND T. ZHA (2010): "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, 77, 665–696.

SIMS, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

STOCK, J. AND M. WATSON (2012): "Disentangling the Channels of the 2007-2009 Recession," *Brookings Papers on Economic Activity*, Spring, 81–135.

STOCK, J. H. AND M. W. WATSON (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15, 101–115.

UHLIG, H. (2005): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381–419.

A Identification in the VAR Toolbox: Technical Details

This Appendix includes additional details on the implementation of the identification schemes discussed in section 7.

A.1 Identification by zero short-run restrictions

The identification by zero short-run restrictions solves the identification problem by setting to zero some of the non-diagonal elements of the structural impact matrix (B), thus reducing the number of unknown coefficients in the B matrix to the same number of equations implied by the condition:

$$\Sigma_u = BB' \tag{A.1}$$

The number of zeros that need to be imposed to achieve identification depend on the number of variables in the VAR – and, crucially, they increase at a faster rate than the numbers of endogenous variables. As discussed in section 7.1, in a simple bivariate VAR a single zero restriction is enough to achieve identification. To see that, consider the system of equations implied by $\Sigma_u = BB'$, namely:

$$\begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ - & \sigma_r^2 \end{bmatrix} = \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}}_{B'} \Rightarrow \begin{cases} \sigma_y^2 = b_{11}^2 + b_{12}^2 \\ \sigma_{yr}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_{yr}^2 = b_{11}b_{21} + b_{12}b_{22} \\ \sigma_r^2 = b_{21}^2 + b_{22}^2 \end{cases} \tag{A.2}$$

The above system has four unknowns but only three independent equations, as the second and the third equation are identical. When setting $b_{12} = 0$, the system of equations (A.2) becomes:

$$\begin{cases} \sigma_y^2 = b_{11}^2, \\ \sigma_{yr}^2 = b_{11}b_{21}, \\ \sigma_r^2 = b_{21}^2 + b_{22}^2. \end{cases} \tag{A.3}$$

which can be easily solved to get:

$$\begin{cases} b_{11} = \sigma_y, \\ b_{21} = \sigma_{yr}^2 / \sigma_y^2, \\ b_{22} = \sqrt{\sigma_r^2 - \frac{\sigma_{yr}^2}{\sigma_y^2}}. \end{cases} \tag{A.4}$$

which shows that the VAR is identified.

This identification scheme is often referred to as 'Cholesky' identification. The reason is the following. A symmetric and positive-definite matrix like Σ_u can always be decomposed as:

$$\Sigma_u = \begin{bmatrix} \sigma_y^2 & \sigma_{yr}^2 \\ \sigma_{yr}^2 & \sigma_r^2 \end{bmatrix} = \begin{bmatrix} p_{11} & 0 \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ 0 & p_{22} \end{bmatrix} = PP' \tag{A.5}$$

where the lower triangular matrix P is known as the Cholesky factor of Σ_u. Recalling that (i) $\Sigma_u = BB'$ and (ii) we assumed that B is also lower triangular (i.e. $b_{21} = 0$), it follows that $P = B$. In other words, instead of solving by hand the system of equations (A.3), we can ask Matlab to compute the Cholesky factor of Σ_u. This is particularly useful for large VARs, since the system of equations implied by $\Sigma_u = BB'$ becomes increasingly complex as the dimensionality of the VAR increases. With reference to the example in section 7.1, the Cholesky decomposition of Σ_u can be implemented by executing the following command in the Matlab command window:

```
>> chol(VAR.sigma,'lower')
       0.5377          0
       0.1454     0.3552
```

which can be shown being the solution to (A.4).

A.2 Identification by zero long-run restrictions

The identification by zero long-run restrictions solves the identification problem by setting to zero some of the non-diagonal elements of the structural long-run matrix:

$$C = (1 - \Phi)^{-1} B, \tag{A.6}$$

thus reducing the number of unknown coefficients in the C matrix to the same number of equations implied by the condition:

$$CC' = \left((I - \Phi)^{-1} \right) \Sigma_u \left((I - \Phi)^{-1} \right)' \tag{A.7}$$

The number of zeros that need to be imposed to achieve identification depend on the number of variables in the VAR – and, as for the case of zero-short run restrictions, they increase at a faster rate than the numbers of endogenous variables. As discussed in section 7.2, in a simple bivariate VAR a single zero restriction is enough to achieve identification.

To see that, define $\Omega \equiv \left((I - \Phi)^{-1} \right) \Sigma_u \left((I - \Phi)^{-1} \right)'$ and note that Ω is a known 2×2 positive-definite symmetric matrix. Thus, Ω admits a unique Cholesky decomposition, given by:

$$\Omega = PP', \tag{A.8}$$

where the lower triangular matrix P is the Cholesky factor of Ω. Because of the assumption that C is lower triangular, it follows that $P = C$. Once C is known, we can recover B from (A.6).

A.3 Identification by sign restrictions

The identification by sign restrictions solves the identification problem by drawing a large number (J) of candidate structural impact matrices (B_j) that satisfy the condition:

$$\Sigma_u = B_j B_j' \qquad j = 0, 1, ..., J \tag{A.9}$$

and retaining those whose elements satisfy a set of *a priori* signs restrictions on a subset or all elements of B_j. Such prior beliefs on the sign of the impact of shocks on the endogenous variables are typically informed by theoretical models. To achieve identification, the sign restrictions need to uniquely identify the shocks of interest.

To generate a candidate structural impact matrix, the VAR Toolbox follows the approach proposed by Rubio-Ramirez et al. (2010):

$$B_j = PQ_j \tag{A.10}$$

where Q_j denotes the orthogonal factor in the QR factorization of a random matrix with elements from the standard normal distribution (j denotes a draw) and P is the Cholesky factor of the reduced-form covariance matrix Σ_u. The matrix B_j is a candidate structural impact matrix because the structural shocks implied by $\varepsilon_{jt} = \left(B_j\right) u_t$ are such that

$$\varepsilon_{jt}\varepsilon_{jt}' = I_2 \tag{A.11}$$

so that condition (A.9) is satisfied.

Differently from the identification schemes described above – where there is a unique B matrix that solves the identification problem – sign restrictions lead to set identification. In other words, the data are potentially consistent with a wide range of B matrices that are all admissible in that they satisfy the sign restrictions.

A.4 Identification with External Instruments

The identification with external instruments solves the identification problem by retrieving one (or more) columns of the structural impact matrix (B) exploiting the information provided by an instrument that is external to the VAR.

In the VAR Toolbox, the identification with external instruments is implemented by a two-stage regression. That is, to isolate the variation in the VAR reduced-form residuals that are due to the structural shock of interest, the VAR Toolbox estimates equations (31) and (32) sequentially. As discussed in Section 7.4, this two-stage approach normalizes the coefficients of the structural impact matrix B by b_{11}. To recover the true values of the elements of B, the VAR Toolbox follows Gertler and Karadi (2015) (footnote 4).

Specifically, the condition $\Sigma_u = BB'$ implies

$$b_{11}^2 = \sigma_y^2 - b_{12}^2 \tag{A.12}$$

as well as

$$b_{11}^2 = \sigma_y^2 - b_{12}^2 \tag{A.13}$$

A.5 Combining sign restrictions and external instruments

The external instruments and sign restriction identification approaches can be combined as proposed by CesBianchiSokol2020. To have a meaningful example we need a trivariate VAR

$$\begin{bmatrix} y_t \\ r_t \\ x_{3t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{12} & b_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \\ \varepsilon_{3t} \end{bmatrix}. \tag{A.14}$$

For simplicity, and without loss of generality, we consider the case where we identify only the first structural shock (labelled ε_t^{IV}) with the external instruments approach. Without loss of generality we choose this shock to be the shock associated with the first equation, so we set $\varepsilon_t^b \equiv \varepsilon_t^{Demand}$. We can then identify the remaining $n-1$ shocks, namely $\varepsilon_t^{B^{SR}} \equiv (\varepsilon_t^{MonPol}, \varepsilon_{3t}')'$, with sign restrictions.

We start by partitioning the matrix B into a column vector B^{IV}, which captures the impact of the shock in the first equation (e_t^b), and a matrix B^{SR}, whose columns capture the impact of the shocks in the remaining 2 equations ($e_t^{B^{SR}}$):

$$B = \begin{bmatrix} B^{IV} & B^{SR} \end{bmatrix}, \tag{A.15}$$

where b is an 3×1 vector and B^{SR} is a 3×2 matrix. Assuming that a valid instrument exists, the first column of the B matrix (b) can be easily identified as explained in the previous section.

We now show how to combine the external instruments identification approach with a standard sign restriction approach to identify the remaining structural shocks ($e_t^{B^{SR}}$) 'conditional' on the shock identified with the external instrument (e_t^b). To identify $\mathcal{B}$ (i.e. the contemporaneous impact of the remaining shocks) we proceed as follows. First,

using (34), we re-write the covariance matrix of the reduced-form residuals as:

$$\Sigma_u = BB' = \left[\begin{array}{cc} b & B^{SR} \end{array} \right] \left[\begin{array}{cc} b & B^{SR} \end{array} \right]'. \tag{A.16}$$

As we have seen above, this decomposition of the covariance matrix is not unique. Let P be the Cholesky decomposition of the covariance matrix Σ_u, and let Q be an orthonormal matrix such that $QQ' = I$. Then we can write:

$$\Sigma_u = PP' = PQQ'C' = (PQ)\,(PQ)' \tag{A.17}$$

Our strategy consists precisely in constructing a large number of orthonormal matrices Q that satisfy the following condition:

$$CQ = \left[\begin{array}{cc} b & B^{SR} \end{array} \right],$$

where b is identified via the external instrument and B^{SR} satisfies a set of sign restrictions. For example, assume we have an instrument for a monetary policy shock and we want to identify the effects of deamnd and supply shocks conditional on the monetary policy shock, namely:

	Monetary Policy	Demand (ε_t^{Demand})	Supply (ε_t^{MonPol})
Policy Rate (y_t)	Proxy	+	
Price (r_t)	Proxy	+	-
Quantity (x_{3t})	Proxy	+	+

We do that in three steps.

1. Find a normal vector q of dimension $n \times 1$ that rotates the first column of C, the Cholesky decomposition of Σ_u, into the vector b. That is, we find a $n \times 1$ normal vector q such that the following equality holds:

$$Cq = b \tag{A.18}$$

2. Given q, build the remaining $n-1$ columns of an orthonormal matrix Q following a standard Gram-Schmidt process.[14] That is, find an $(n \times n-1)$ matrix $\mathcal{Q}$ such that the following equality holds:

$$\begin{bmatrix} q & \mathcal{Q} \end{bmatrix} \begin{bmatrix} q & \mathcal{Q} \end{bmatrix}' = QQ' = I. \tag{A.19}$$

The matrix CQ then represents a candidate identification scheme because:

$$CQ = C \begin{bmatrix} q & \mathcal{Q} \end{bmatrix} = \begin{bmatrix} b & B^{SR} \end{bmatrix} = B. \tag{A.20}$$

3. Check that B^{SR} satisfies our set of sign restrictions. If it does, we retain the

[14] Let j index the columns of Q. Let Q_{j-1} denote the first $j-1$ columns of Q, such that $Q_{2-1} = Q_1 = q_1$. Let x_j be a draw from a Normal distribution on $\mathbb{R}^N$. Then the $j-$th column of Q can be constructed as:

$$q_j = \frac{\left(I_N - Q_{j-1}Q_{j-1}'\right) x_j}{\left\| \left(I_N - Q_{j-1}Q_{j-1}'\right) x_j \right\|}.$$

matrix Q. If does not, we repeat steps (1) and (2) until we obtain a matrix B^{SR} that satisfies the restrictions.

Finally we repeat steps (1)-(2)-(3) until we have M matrices B_i (with $i = 1, 2, ..., M$) consistent with our identification restrictions. This completes the (set) identification of structural matrix B.

B Structural Dynamic Analysis

B.1 Impulse response functions

Impulse response functions (IR) allow us to answer the following question: 'What is the response over time of each of the variables in a VAR to an increase in the current value of one of the structural innovations, assuming that (i) the structural innovation returns to zero in subsequent periods and (ii) all other structrual innovations are equal to zero?

Of course, the implied thought experiment of shocking the innovations of one equation while holding the others constant makes sense only when the innovations are uncorrelated across equations – which can be done only once we know the structural representation of the VAR, i.e. once we have identified the B matrix.

To show how to compute impulse response functions, consider our simple bivariate VAR in its structural representation

$$\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{Demand} \\ \varepsilon_t^{MonPol} \end{bmatrix}, \tag{B.1}$$

Then, define a 2×1 vector of impulse selection (s) that take value of one for the strcuctural shock that we wan to consider. For example, to compute the IR to the first structural shocks we define s as:

$$s = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The impulse responses to the structural shock ε_t^{Demand} can be easily computed with the following equation

$$x_t = \Phi x_{t-1} + B s_t,$$

which can be computed recursively as follows

$$\begin{cases} IR_h = Bs, & \text{for } h = 1, \\ IR_h = \Phi \cdot IR_{h-1} & \text{for } h = 2, ..., H. \end{cases}$$

B.2 Forecast error variance decompositions

Forecast error variance decompositions (VD) answer the following question: What portion of the variance of the forecast error that is made when predicting $x_{i,t+h}$ is due to each structural shock in ε_t (on average, over the sample period)? As such, VD provide information about the relative importance of each structural shock in affecting the variables in the VAR

To show how to compute forecast error variance decompositions, consider the 1-step ahead forecast error in our simple bivariate VAR:

$$x_{t+1} - \mathbb{E}\left[x_{t+1}\right] = x_{t+1} - \mathbb{E}\left[\Phi x_t + u_{t+1}\right] = x_{t+1} - \Phi x_t = u_{t+1}$$

where note that the 1-step ahead forecast error is the 1-step ahead reduced form residual. As we know from the previous setion, the reduced-form residual is related to the structural shocks throught the following equations

$$\begin{cases} u_{1t+1} = b_{11}\varepsilon_{1t+1} + b_{12}\varepsilon_{2t+1}, \\ u_{2t+1} = b_{21}\varepsilon_{1t+1} + b_{22}\varepsilon_{2t+1}. \end{cases}$$

So, what is the variance of the forecast error?

$$\mathbb{V}\left(u_{yt}\right) = b_{11}^2 \mathbb{V}\left(\varepsilon_{1,t+1}\right) + b_{12}^2 \mathbb{V}\left(\varepsilon_{2,t+1}\right) = b_{11}^2 + b_{12}^2$$
$$\mathbb{V}\left(u_{rt}\right) = b_{21}^2 \mathbb{V}\left(\varepsilon_{1,t+1}\right) + b_{22}^2 \mathbb{V}\left(\varepsilon_{2,t+1}\right) = b_{21}^2 + b_{22}^2$$

whihc follows from the fact that the variacne of ε_t is 1 and the structural shocks are orthogonal to each other. The final step to compute VD is to ask: what portion of the variance of the forecast error is due to each structural shock? The answer is given by teh following equation

$$\underbrace{\begin{cases} VD_y^{\varepsilon_1} = \frac{b_{11}^2}{b_{11}^2+b_{12}^2} \\ VD_y^{\varepsilon_2} = \frac{b_{12}^2}{b_{11}^2+b_{12}^2} \end{cases}}_{\text{This sums up to 1}} \qquad \underbrace{\begin{cases} VD_r^{\varepsilon_1} = \frac{b_{21}^2}{b_{21}^2+b_{22}^2} \\ VD_r^{\varepsilon_2} = \frac{b_{22}^2}{b_{21}^2+b_{22}^2} \end{cases}}_{\text{This sums up to 1}}$$

B.3 Historical decompositions

Historical decompositions (HD) answer the following question: What portion of the deviation of $x_{i,t}$ from its unconditional mean is due to each structural shock ε_t?

We showed before that, in the absence of shocks, the variables of a (stable) VAR will converge to their unconditional mean (i.e. their equilibrium values or steady state). As we ahev seen with impulse responses, when a structural shock hits, the endogenous variables move away from their equilibrium and then only slowly go back to it. If another shock hits in the next period, the variables will now be away from equilibrium because of (i) the effect of the new shock and (ii) the persistent effect of the old shock. Historical decompositions allow us to know, at each point in time, what shock is responsible for keeping the endogenous variables away from their steady state.

To show how to compute historical decompositions, we start from the Wold decomposition of a VAR in Equation (**??**), according to which each observation can be re-written as the cumulative sum of the structural shocks. In particular, we consider the Wold representation of our simple bivariate VAR for $t = 2$:

$$r = \underbrace{\Phi^2 x_0}_{init_2} + \underbrace{\Phi B}_{\mathbf{\Theta_1}} \varepsilon_1 + \underbrace{B}_{\mathbf{\Theta}_2} \varepsilon_2,$$

which allows us to write r as a function of present and past structural shocks (ε_1 and ε_2) and the initial condition (x_0). In matrix form:

$$\begin{bmatrix} x_{1,2} \\ x_{2,2} \end{bmatrix} = \begin{bmatrix} init_{y,2} \\ init_{r,2} \end{bmatrix} + \begin{bmatrix} \theta_{11}^1 & \theta_{12}^1 \\ \theta_{21}^1 & \theta_{22}^1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,1} \\ \varepsilon_{2,1} \end{bmatrix} + \begin{bmatrix} \theta_{11}^2 & \theta_{12}^2 \\ \theta_{21}^2 & \theta_{22}^2 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,2} \\ \varepsilon_{2,2} \end{bmatrix}$$

Therefore r can be expressed as

$$\begin{cases} x_{1,2} = init_{y,2} + \theta_{11}^1 \varepsilon_{1,1} + \theta_{12}^1 \varepsilon_{2,1} + \theta_{11}^2 \varepsilon_{1,2} + \theta_{12}^2 \varepsilon_{2,2} \\ x_{2,2} = init_{r,2} + \theta_{21}^1 \varepsilon_{1,1} + \theta_{22}^1 \varepsilon_{2,1} + \theta_{21}^2 \varepsilon_{1,2} + \theta_{22}^2 \varepsilon_{2,2} \end{cases}$$

The historical decomposition is given by

$$\underbrace{\begin{cases} HD_{y,2}^{\varepsilon_1} = \theta_{11}^1 \varepsilon_{1,1} + \theta_{11}^2 \varepsilon_{1,2} \\ HD_{y,2}^{\varepsilon_2} = \theta_{12}^1 \varepsilon_{2,1} + \theta_{12}^2 \varepsilon_{2,2} \\ HD_{y,2}^{init} = init_{y,2} \end{cases}}_{\text{This sums up to } x_{1,2}} \quad \underbrace{\begin{cases} HD_{r,2}^{\varepsilon_1} = \theta_{21}^1 \varepsilon_{1,2} + \theta_{21}^0 \varepsilon_{1,2} \\ HD_{r,2}^{\varepsilon_2} = \theta_{22}^1 \varepsilon_{2,2} + \theta_{22}^0 \varepsilon_{2,2} \\ HD_{r,2}^{init} = init_{r,2} \end{cases}}_{\text{This sums up to } x_{2,2}}$$

where $HD_{y,2}^{\varepsilon_1}$ is the contribution of present and past shocks to ε_1 to y in period $t = 2$. Note that y, for example, is equal to the sum of the contribution of (i) present and past shocks to ε_1, (ii) present and past shocks to ε_2, and (iii) the initial condition. The first two elements are obvious: if shocks are persistent we would expect today's calue of y to be affectd by present and recent shocks. The third element depends on how far the first observation in our data (x_0) is from its unconditinoal mean. In this example, we are assuming that the unconditional mean of the data is zero. So if x_0 is very different from 0, the the initial condition will matter for many periods until it will become asymptotically small.