Financial Econometrics II

Block 1: Forecasting and simulations with ARMA/VAR/SVAR

Michał Rubaszek
SGH Warsaw School of Economics
TOPICS

1. Introduction to R

2. ARMA models

3. VAR models
Meeting 1. Introduction to R
Content of R codes

1. Operations on vectors and matrices

2. Conditioning, loops, defining functions

3. Importing data (read.csv, Quandl, quantmod, Eurostat)

4. Converting and plotting data (ts, zoo, xts)

5. Simple vs. compound interest rate
Rates of return / growth rates

Simple rate of return:

\[ Y_t = (1 + R_t)Y_{t-1} \iff R_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}} \]

Compound interest rate (\(m\) is compounding frequency):

\[ Y_t = \left(1 + \frac{R_{m,t}}{m}\right)^m Y_{t-1} \]

Continuously compound interest rate:

\[ Y_t = \lim_{m \to \infty} \left(1 + \frac{R_{m,t}}{m}\right)^m Y_{t-1} = \exp(r_t) Y_{t-1} \]

Logarithmic rate of return:

\[ Y_t = \exp(r_t)Y_{t-1} \iff r_t = \ln(Y_t/Y_{t-1}) \]

Notice: \(1 + R = \exp(r) \iff r = \ln(1 + R)\)
Simple returns:

- Easy to calculate for a portfolio of assets: 
  \[ R_p = \sum_{k=1}^{K} w_k R_k \]
- Easy to communicate to non-statisticians
- Not symmetric nor additive...

Log returns:

- Symmetric and additive
- Easy to communicate to statisticians
- Difficult to calculate for a portfolio of assets: 
  \[ r_p \neq \sum_{k=1}^{K} w_k r_k \]

We will work with log returns
Exercises

Exercise 1.1.
Write an algorithm, which would allow to calculate the roots of the equation:

\[ e^x - (x + 1)^2 = 0 \]

knowing that they are in the interval \((-3, 3)\).

[Hint: make two loops with functions for and while]

Exercise 1.2.
Create a function \( invVal(Y, h, R, m) \) that will calculate the value of investment \( Y \) after \( h \) years, given that the annual interest rate is \( R \) and compound frequency \( m \).

Use the function to calculate the value of 1000PLN after 1 year for \( m = \{1, 2, 4, \infty\} \) and \( R_m = 10\% \).
Exercise 1.3.
Using the `eurostat` package import to R the annual growth rate of real GDP in Poland (at quarterly frequency). Write a series as a `zoo` object and make a plot. What was the average growth rate over the last 10 years?

Exercise 1.4.
Import daily data for the WIG index from the Internet to R. After converting the series to a `zoo` object, make a panel of figures for
- historic prices
- logarithmic growth rates
- ACF for levels
- ACF for growth rates
Meeting 2. ARMA models
Plan for today

1. Calculating impulse-response functions
2. Testing for unit root
3. Estimating ARMA model
4. Forecasting with ARMA model
IRF – impulse response function

Impulse response function – IRF:
describe how variable $y_t$ reacts over time to exogenous impulse $\epsilon_t$.

Moving Average model:

$$y_t = \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots$$

Formula for IRF:

$$IRF_k = \theta_k = \frac{\partial y_t}{\partial \epsilon_{t-k}} = \frac{\partial y_{t+k}}{\partial \epsilon_t}$$

How to calculate IRF for a model?
Transform model to a moving average (MA) form
Calculating IRF: AR(1) model

**AR model:**

\[ y_t = \rho y_{t-1} + \epsilon_t \quad \leftrightarrow \quad (1 - \rho L)y_t = \epsilon_t \]

\[ y_t = (1 - \rho L)^{-1} \epsilon_t = \sum_{k=0}^{\infty} \rho^k \epsilon_{t-k} \quad \text{[} + \lim_{s \to \infty} \rho^s E(y_{t-s}) \text{]} \]

\[ \theta_k = \rho^k \]

**AR model with a constant:**

\[ y_t = \alpha + \rho y_{t-1} + \epsilon_t \]

\[ y_t = (1 - \rho L)^{-1} (\epsilon_t + \alpha) = \frac{\alpha}{1 - \rho} + \sum_{k=0}^{\infty} \rho^k \epsilon_{t-k} \]

\[ E(y_t) = \mu = \frac{\alpha}{1 - \rho} \quad \text{and} \quad \theta_k = \rho^k \]
Calculating IRF: AR(P) model

**AR(P) model:**

\[ y_t = \sum_{p=1}^{P} \rho_p y_{t-p} + \epsilon_t \]

\[ \rho(L)y_t = \prod_{p=1}^{P} (1 - \lambda_p L)y_t = \epsilon_t \]

\[ y_t = \prod_{p=1}^{P} (1 - \lambda_p L)^{-1} \epsilon_t \]

AR(P) as a multiplication of \( P \) AR(1) processes

**AR(2) case:**

\[ y_t = (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \epsilon_t = (1 - \lambda_1 L)^{-1} z_t = \sum_{k=0}^{\infty} \lambda_1^k z_{t-k} \]

\[ z_t = (1 - \lambda_2 L)^{-1} \epsilon_t = \sum_{k=0}^{\infty} \lambda_2^k \epsilon_{t-k} \]
Exercises

Exercise 2.1.
Write the IRF $y_{t+k}$ with respect to $\epsilon_t$ for the following processes:

A. $y_t = 0.8y_{t-1} + \epsilon_t$

B. $y_t = 2 + 0.5y_{t-1} + \epsilon_t$

C. $y_t = 1.3y_{t-1} - 0.4y_{t-2} + \epsilon_t$

D. $y_t = 0.8y_{t-1} + \epsilon_t - 0.5\epsilon_{t-1}$

E. $y_t = y_{t-1} + \epsilon_t$
Unit root

For AR model: \( y_t = \rho y_{t-1} + \epsilon_t \)
\( y_t = \sum_{k=0}^{\infty} \rho^k \epsilon_{t-k} \) \( + \lim_{s \to \infty} \rho^s E(y_{t-s}) \)

If \( \rho = 1 \) then

- \( \lim_{s \to \infty} \rho^s E(y_{t-s}) \neq 0 \)
- \( \lim_{s \to \infty} \rho^s = 1 \)

This means that the impact of a shock is not decaying and that the process is not returning to an equilibrium value. It is non-stationary

For \( \rho < 1 \) we might calculate so-called half-life:

\( HL = \frac{\ln 0.5}{\ln \rho} \leftrightarrow \rho^{HL} = 0.5 \)
**Unit root: stationarity**

**Definition of weak stationarity:**
A process is said to be covariance-stationary, or weakly stationary, if its first and second (unconditional) moments are time invariant, and for each period $t$ are equal to:

\[
\begin{align*}
E(Y_t) &= \mu \\
Var(Y_t) &= \gamma_0 = \sigma^2 \\
Cov(Y_t, Y_{t-s}) &= \gamma_s
\end{align*}
\]

**Important:**
- unconditional and conditional moments might differ
- Stationarity can be interpreted as „mean reversion” , i.e. that a series should fluctuate around $\mu$ and its volatility around $\sigma$
**Unit root: stationarity**

**Definition of strong stationarity:**

The joint distribution of $Y_{t_1}, Y_{t_2}, ..., Y_{t_s}$ is the same as the joint distribution of $Y_{t_1+k}, Y_{t_2+k}, ..., Y_{t_s+k}$:

$$f(Y_{t_1}, Y_{t_2}, ..., Y_{t_s}) = f(Y_{t_1+k}, Y_{t_2+k}, ..., Y_{t_s+k})$$

- Is not limited to the first two moments
- Implies weak stationarity
- Not particularly useful in practical applications as it cannot be tested...
Augmented Dickey-Fuller test:

\[ \Delta y_t = [\alpha_0 + \alpha_1] + \delta y_{t-1} + \sum_{p=1}^{P} \gamma_p \Delta y_{t-p} + \epsilon_t \]

\( H0: \delta = 0, \) i.e. non-stationarity

\( H1: \delta < 0, \) i.e. stationarity

\[ DF_{cal} = \frac{\hat{\delta}}{S_\delta} \sim DF \]

**Note:** Adding lags of \( \Delta y \) is a parametric correction for possible autocorrelation of the error term \( \epsilon_t \)
**Unit root: tests**

**Phillips-Perron test:**

\[ y_t = [\alpha_0 + \alpha_1] + \rho y_{t-1} + \epsilon_t \]

*H0:* \( \rho = 1 \), i.e. non-stationarity

*H1:* \( \rho < 1 \), i.e. stationarity

\[
PP_{cal} = \left( \frac{\hat{\gamma}_0}{\hat{\gamma}_\infty} \right)^{0.5} \frac{(\hat{\rho} - 1)}{S_\rho} - \frac{T}{2} (\hat{\gamma}_\infty - \hat{\gamma}_0) \left( \frac{S_\rho}{\sqrt{\hat{\gamma}_\infty \hat{\gamma}_0}} \right) \sim PP
\]

where \( \hat{\gamma}_0 \) and \( \hat{\gamma}_\infty \) are variance and long-term variance for residuals \( \epsilon_t \).

**Note:** If \( \hat{\gamma}_0 = \hat{\gamma}_\infty \) then \( PP_{cal} = DF_{cal} \). In other case we have non-parametric correction for possible autocorrelation of the error term \( \epsilon_t \)
**Unit root: tests**

**KPSS test:**

\[ y_t = x_t + z_t \]

\[ x_t = x_{t-1} + \nu_t, \nu_t \sim WN(0, \sigma_v^2) \]

\[ z_t = [\alpha_0 + \alpha_1] + \epsilon_t \]

\( H0: \sigma_v^2 = 0, \) i.e. stationarity

\( H1: \sigma_v^2 > 0, \) i.e. non-stationarity

\[ \text{KPSS}_{cal} = \frac{1}{T^2} \sum_{t=1}^{T} S_t^2 \]

\[ \sim \text{KPSS} \]

where \( \hat{\gamma}_\infty \) is the long-run variance of residuals \( \hat{\epsilon}_t \) from regression of \( y_t \) on a constant and a trend (depending on a specification) and \( S_t = \sum_{s=1}^{t} \hat{\epsilon}_s \)
Unit root: tests

Important:
For persistent processes and small samples the power of ADF and PP tests is low, whereas the KPSS test is subject to size distortion
[illustration in the Monte Carlo example in the R file]

Implication:
- Be careful while differentiating the data
- Economic knowledge might be better advice that the tests
Exercise 2.2.
Import data for the US economy over the years 1860-1970 with commands:
> require(urca)
> data(nporg)
For each series decide whether to use logs or not. Test for stationarity.
What are the economic reasons of non-stationarity?

Exercise 2.3.
Import data for HICP YoY inflation for a selected EU country with the eurostat package. Decide on the level of integration of the downloaded variable using ADF, PP and KPSS tests.
Specification of ARMA(P,Q) model:

\[ y_t = [\alpha_0 + \alpha_1 t] + \sum_{p=1}^{P} \rho_p y_{t-p} + \sum_{q=0}^{Q} \gamma_q \varepsilon_{t-q} \]

\[ \rho(L)y_t = [\alpha_0 + \alpha_1 t] + \gamma(L)\varepsilon_t \]

Equilibrium value for stationary ARMA model:

\[ E(y_t) = \mu = \frac{\alpha_0}{1-\sum_{p=1}^{P} \rho_p} \]

Specification of ARIMA(P,D,Q) model:

\[ \rho(L)(1 - L)^D y_t = [\alpha_0 + \alpha_1 t] + \gamma(L)\varepsilon_t \]
ARMA model

Why do we need ARMA models?

- For analysing the properties of univariate time series
- Seasonal adjustment
Let us consider AR(1):

\[ y_t = \alpha + \rho y_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2_\epsilon) \]

**Conditional likelihood of a single observation:**

\[
p(y_t | \alpha, \rho, y_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2_\epsilon}} \exp \left( -\frac{(y_t - \alpha - \rho y_{t-1})^2}{2\sigma^2_\epsilon} \right)
\]

**Likelihood of all observations:**

\[
p(y_1, y_2, ..., y_T | \alpha, \rho) = p(y_1 | \alpha, \rho) \times p(y_2 | \alpha, \rho, y_1) \times ... \times p(y_T | \alpha, \rho, y_{T-1})
\]

where \( p(y_1 | \alpha, \rho) = \frac{1}{\sqrt{2\pi\sigma^2_\epsilon}} \exp \left( -\frac{(y_1 - \alpha)^2}{2\sigma^2_\epsilon} \right) \)

- If we neglect \( p(y_1 | \alpha, \rho) \) : conditional ML estimator (= OLS estimator)
- If we include \( p(y_1 | \alpha, \rho) \) : full ML estimator
  (the derivation of ML function for ARMA(P,Q) with the Kalman filter - link)
Information criteria:

Akaike: \[ AIC = -2 \frac{\ell}{T} + 2 \frac{K}{T} \]

Schwarz: \[ BIC = -2 \frac{\ell}{T} + 2 \frac{K}{T} \ln(T) \]

Hannan-Quinn: \[ HQIC = -2 \frac{\ell}{T} + 2 \frac{K}{T} \ln(\ln T) \]

where \( K \) is the number of estimated parameters and \( \ell = \ln \mathcal{L} \) is the log-likelihood. We choose the model with the lowest IC.

Notice:
- \( K(SIC) \leq K(HQIC) \leq K(AIC) \)
- IC depends on the fit (log-likelihood) and penalty on the number of params.
Likelihood ratio test:

H0: the fit of big ARMA (m params more) is the same as fit of small ARMA

\[ LR = -2(\ell_r - \ell_u) \sim \chi^2(m) \]

where \( m \) is the number of additional params. and \( \ell \) is the log-likelihood for restricted (small) and unrestricted (big) models.

Autocorrelation (portmanteau) Ljung-Box test:

H0: \( \rho_{e,j} = 0 \) for \( j = 1,2, ..., J \)

\[ LB = T^2 \sum_{j=1}^{J} \frac{1}{T-j} \hat{\rho}_{e,j}^2 \sim \chi^2(J - (P + Q)) \]
ARMA model - forecasting

ARMA(P,Q):
\[ y_t = [\alpha_0 + \alpha_1 t] + \sum_{p=1}^{P} \rho_p y_{t-p} + \sum_{q=0}^{Q} \gamma_q \epsilon_{t-q} \]

MA representation:
\[ y_t = \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots \]

Point forecast (calculated recursively):
\[ y_{T+h}^f = [\alpha_0 + \alpha_1 t] + \sum_{p=1}^{P} \rho_p y_{T+h-p}^f + \sum_{q=0}^{Q} \gamma_q \epsilon_{T+h-q}^f \]
where \( \epsilon_{T+h}^f = 0 \) for \( h > 0 \) and \( y_{T+h}^f = y_{T+h} \) for \( h \leq 0 \)

Forecast error:
\[ y_{T+H} - y_{T+H}^f = \sum_{h=0}^{H-1} \theta_h \epsilon_{T+H-h} \]

Forecast variance (only due to stochastic term):
\[ Var(y_{T+H}) = \sum_{h=0}^{H-1} \theta_h^2 \]
Exercises

Exercise 2.4.
For the ARMA(2,0) model:

\[ y_t = 1.1y_{t-1} - 0.3y_{t-2} + \epsilon_t, \epsilon_t \sim N(0,1) \]

- check if the model is stationary
- Knowing \( y_T = 1 \) and \( y_{T-1} = 2 \) calculate the forecast for periods \( T + 1 \) and \( T + 2 \).
- Write the model in \( MA(\infty) \) form. Calculate the values for the first three coefs.
- Calculate point and 95% interval forecast for \( T + 2 \)

Exercise 2.5.
Import data for HICP YoY with the \texttt{eurostat} package for a selected EU country and:

- Choose the specification of the ARMA model with the Schwarz criterion
- Convert the model to \( MA(\infty) \) form
- Verify the model for autocorrelation
- Calculate point and density forecasts
Meeting 3. VAR models
Plan for today

1. Estimating VAR model
2. Structural of VAR
3. Impulse-response function (IRF)
4. Forecast error variance decomposition
5. Historical decomposition
6. Forecasting with VAR model
Vector autoregression model - VAR

- VAR is an extension of univariate autoregression model to multivariate time series data, in which all variables are treated as endogenous (Sims critique).

- VAR is a useful model for analysing the dynamic behaviour of economic and financial time series. It is also a standard tool used for forecasting.

- Structural VAR as an important „story-telling” model.
Specification of a VAR model

VAR(p) model:

\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \epsilon_t , \epsilon_t \sim N(0, \Sigma) \]

where:

- \( y_t = [y_{1t} \ y_{2t} \ \cdots \ y_{nt}]' \) vector of \( n \) endogenous variables
- \( \epsilon_t = [\epsilon_{1t} \ \epsilon_{2t} \ \cdots \ \epsilon_{nt}]' \) vector of error terms
- \( C = [c_1 \ c_2 \ \cdots \ c_n]' \) vector of constants
- \( A_p = [a_{ij,p}]_{n \times n} \) matrix of parameters
- \( \Sigma = [\sigma_{ij}]_{n \times n} \) covariance matrix
IRF: VAR(1) case

VAR model:
\[ y_t = Ay_{t-1} + \epsilon_t \]
\[ y_t = \epsilon_t + A\epsilon_{t-1} + A^2\epsilon_{t-2} + A^3\epsilon_{t-3} + \ldots \]

VMA representation:
\[ y_t = \Psi_0\epsilon_t + \Psi_1\epsilon_{t-1} + \Psi_2\epsilon_{t-2} + \Psi_3\epsilon_{t-3} + \ldots \]

\[ \Psi_k = \left[ \psi_{ij,k} \right]_{n \times n} = \frac{\partial y_t}{\partial \epsilon_{t-k}} = \frac{\partial y_{t+k}}{\partial \epsilon_t} : \text{ } n \times n \text{ matrix if IRFs} \]

\[ \Psi_k = A^k \]
Stationarity: VAR(1) case

\[ y_t = \Psi_0 \epsilon_t + \Psi_1 \epsilon_{t-1} + \Psi_2 \epsilon_{t-2} + \Psi_3 \epsilon_{t-3} + \cdots \]

VAR is stationary if:

\[ \lim_{k \to \infty} \Psi_k = \lim_{k \to \infty} A^k = 0 \]

**Spectral decomposition:**

\[ A = V \Lambda V^{-1} \]

\( \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \)  eigenvalues

\( V = [v_1 \ v_2 \ \ldots \ v_n] \)  eigenvectors

\[ y_t = V \Lambda V^{-1} y_{t-1} + \epsilon_t \quad \leftrightarrow \quad \tilde{y}_t = \Lambda \tilde{y}_{t-1} + \tilde{\epsilon}_t \]

where \( \tilde{y}_t = V^{-1} y_t \) and \( \tilde{\epsilon}_t = V^{-1} \epsilon_t \)
Stationarity: VAR(1) case

\[ \tilde{y}_t = \Lambda \tilde{y}_{t-1} + \tilde{\epsilon}_t \]

\[
\begin{bmatrix}
\tilde{y}_{1t} \\
\tilde{y}_{2t} \\
\vdots \\
\tilde{y}_{nt}
\end{bmatrix} =
\begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{bmatrix}
\begin{bmatrix}
\tilde{y}_{1t-1} \\
\tilde{y}_{2t-1} \\
\vdots \\
\tilde{y}_{nt-1}
\end{bmatrix} +
\begin{bmatrix}
\tilde{\epsilon}_{1t} \\
\tilde{\epsilon}_{2t} \\
\vdots \\
\tilde{\epsilon}_{3t}
\end{bmatrix}
\]

We have \( n \) univariate AR(1) processes, hence \( \tilde{y}_t \) is stationary of all characteristic roots are lower than unity \( |\lambda_i| < 1 \).

\[ y_t = V\tilde{y}_t: \text{ our observables as a linear combination of } n \text{ AR(1) processes} \]

\[ \lim_{k \to \infty} \Psi_k = \lim_{k \to \infty} A^k = \lim_{k \to \infty} V\Lambda^k V^{-1} = 0 \text{ only if } |\lambda_i| < 1 \text{ for } i=1,2, \ldots, n \]
Stationarity: VAR(P) case

VAR(P) model:
\[ y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \epsilon_t \]

Canonical form of VAR(p):
\[ y_t^* = A^* y_{t-1}^* + \epsilon_t^* \]

\[ A^* = \begin{bmatrix} A_1 & A_2 & \cdots & A_p \\
I & 0 & \cdots & \cdots \\
0 & I & \cdots & 0 \\
0 & 0 & \cdots & 0 \end{bmatrix}, y_t^* = \begin{bmatrix} y_t \\
y_{t-1} \\
\cdots \\
y_{t-p+1} \end{bmatrix}, \epsilon_t^* = \begin{bmatrix} \epsilon_t \\
0 \\
\cdots \\
0 \end{bmatrix} \]

- Model is stationary if the roots of characteristic polynomial for \( A^* \) are \(|\lambda_i^*| < 1\)
- VMA form can be calculated for canonical VAR

\[ y_t^* = \Psi_0^* \epsilon_t^* + \Psi_1^* \epsilon_{t-1}^* + \Psi_2^* \epsilon_{t-2}^* + \Psi_3^* \epsilon_{t-3}^* + \cdots \]

- \( \Psi_k \) is the \( n \times n \) upper part of \( \Psi_k^* \)
Estimating the VAR model

\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma) \]

\[ y'_t = C' + y'_{t-1}A'_1 + y'_{t-2}A'_2 + \cdots + y'_{t-p}A'_p + \epsilon'_t \]

\[ y'_t = x'_t B + \epsilon'_t \]

\[
\begin{bmatrix}
1 \\
y_{t-1} \\
y_{t-2} \\
\vdots \\
y_{t-p}
\end{bmatrix}, \quad B = 
\begin{bmatrix}
C' \\
A'_1 \\
A'_2 \\
\vdots \\
A'_p
\end{bmatrix}
\]

\[ Y = XB + \varepsilon \]

\[
\begin{bmatrix}
y'_1 \\
y'_2 \\
y'_3 \\
\vdots \\
y'_{T}
\end{bmatrix}, \quad X = 
\begin{bmatrix}
x'_1 \\
x'_2 \\
x'_3 \\
\vdots \\
x'_{T}
\end{bmatrix}, \quad \varepsilon = 
\begin{bmatrix}
\epsilon'_1 \\
\epsilon'_2 \\
\epsilon'_3 \\
\vdots \\
\epsilon'_{T}
\end{bmatrix}
\]
Estimating the VAR model

VAR(p) in matrix notation:
\[ Y = XB + \varepsilon \]

LS estimates
\[ \hat{B} = (X'X)^{-1}XY \]

Residuals
\[ \hat{\varepsilon} = Y - X\hat{B} \]

Estimate of the covariance matrix
\[ \hat{\Sigma} = (T - k)^{-1}(\hat{\varepsilon}'\hat{\varepsilon}) \]
Where \( k = 1 + np \) is the number of parameters in each equations

More details: see p. 16-18 of Dieppe et al. (2016) - link
**VAR model - specification**

### Information criteria:

- **Akaike:**
  \[
  AIC = -2 \frac{\ell}{T} + 2 \frac{K}{T}
  \]

- **Schwarz:**
  \[
  BIC = -2 \frac{\ell}{T} + 2 \frac{K}{T} \ln(T)
  \]

- **Hannan-Quinn:**
  \[
  HQIC = -2 \frac{\ell}{T} + 2 \frac{K}{T} \ln(\ln T)
  \]

where \(K = n(1 + np)\) is the number of parameters and \(\ell = \ln L\) is the log-likelihood.

### Ljung-Box (adjusted portmanteau) autocorrelation test:

\[
LB_{cal} = T^2 \sum_{j=1}^{J} \frac{1}{T-j} \text{tr}(\hat{\Gamma}_j' \hat{\Gamma}_0^{-1} \hat{\Gamma}_j \hat{\Gamma}_0^{-1}) \sim \chi^2(n^2(J-p))
\]

where \(\hat{\Gamma}_j = \frac{1}{T} \sum \epsilon_t \epsilon_{t-j}'\).
**SVAR: structural VAR**

**VAR(P) model, in which shocks have no economic interpretation:**
\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t, \; \varepsilon_t \sim N(0, \Sigma) \]

**SVAR(p) model, in which shocks have interpretation**
\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + D \eta_t, \; \eta_t \sim N(0, I) \]

[equivalent notation, by multiplying both sides by \( D^{-1} \)]
\[ D_0 y_t = F + D_1 y_{t-1} + D_2 y_{t-2} + \cdots + D_p y_{t-p} + \eta_t, \; \eta_t \sim N(0, I) \]

**SVMA representation**
\[ y_t = \mu + \psi_0 \eta_t + \psi_1 \eta_{t-1} + \psi_2 \eta_{t-2} + \psi_3 \eta_{t-3} \cdots \]
where \( \psi_0 = D \)
**SVAR: identification of shocks**

**VAR(p) model:**
\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \epsilon_t, \epsilon_t \sim N(0, \Sigma) \]

**SVAR(p) model:**
\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + D\eta_t, \eta_t \sim N(0, I) \]

**SVMA representation**
\[ y_t = \mu + \psi_0 \eta_t + \psi_1 \eta_{t-1} + \psi_2 \eta_{t-2} + \psi_3 \eta_{t-3} \cdots \]

**How to find matrix \(D\)?**
We need to impose \(\frac{n(n-1)}{2}\) restrictions, taking into account that \(DD' = \Sigma\)

**Short-term restrictions / Cholesky identification:**
We assume that \(D\) is lower triangular matrix

**Long-run restrictions / Blanchard-Quah identification:**
We impose restrictions on the matrix of long-term response \(\psi = \sum_{i=0}^{\infty} \psi_i = (I - A_1 - \cdots - A_p)^{-1} D\)
FEVD: forecast error variance decomposition

SVMA representation

\[ y_t = \mu + \psi_0 \eta_t + \psi_1 \eta_{t-1} + \psi_2 \eta_{t-2} + \psi_3 \eta_{t-3} \ldots \]

Error of forecast for horizon \( h \) due to future shocks

\[ y_{t+h} - y_{t+h|t} = \psi_0 \eta_{t+h} + \psi_1 \eta_{t+h-1} + \cdots + \psi_{h-1} \eta_{t+1} \]

Variance of forecast error for horizon \( h \) due to future shocks

\[ \text{Var}_t(y_{t+h}) = \psi_0 \psi'_0 + \psi_1 \psi'_1 + \cdots + \psi_{h-1} \psi'_{h-1} \]

More details: see p. 101-103 of Dieppe et al. (2016) - [link](#)
FEVD: forecast error variance decomposition

Variance of forecast error for horizon $h$ due to future shocks

$$Var_T(y_{T+h}) = \psi_0 \psi'_0 + \psi_1 \psi'_1 + \ldots + \psi_{h-1} \psi'_{h-1}$$

$$\psi_k = \begin{bmatrix} \psi_{s,11} & \ldots & \psi_{s,1n} \\ \vdots & \ddots & \vdots \\ \psi_{s,n1} & \ldots & \psi_{s,nn} \end{bmatrix}, \psi_s \psi'_s = \begin{bmatrix} \psi_{s,11}^2 + \psi_{s,12}^2 + \ldots + \psi_{s,1n}^2 \\ \vdots \\ \psi_{s,n1}^2 + \psi_{s,n2}^2 + \ldots + \psi_{s,nn}^2 \end{bmatrix}$$

Substituting yields:

$$\sigma_i^2(h) = Var_T(y_{i,T+h}) = \sum_{s=0}^{h-1} (\psi_{s,i1}^2 + \psi_{s,i2}^2 + \ldots + \psi_{s,in}^2) = \sum_{j=1}^{n} (\psi_{0,ij}^2 + \psi_{1,ij}^2 + \ldots + \psi_{h-1,ij}^2)$$

**Contribution of shocks $\eta_{j,T+s}$ to total forecast error variance:**

$$\sigma_{ij}^2(h) = \psi_{0,ij}^2 + \psi_{1,ij}^2 + \ldots + \psi_{h-1,ij}^2$$

so that:

$$\sigma_i^2(h) = \sum_{j=1}^{n} \sigma_{ij}^2(h)$$
**Historical decomposition**

SVMA representation

\[ y_t = \mu + \psi_0 \eta_t + \psi_1 \eta_{t-1} + \psi_2 \eta_{t-2} + \psi_3 \eta_{t-3} \ldots \]

For a single variable

\[ y_{it} = \mu_i + \sum_{j=1}^n (\psi_{0,ij} \eta_{j,t} + \psi_{1,ij} \eta_{j,t-1} + \psi_{2,ij} \eta_{j,t-2} + \psi_{3,ij} \eta_{j,t-3} \ldots) \]

Contribution of shocks \( \eta_{j,t-s} \) the value of \( y_{it} \):

\[ y_{it,j} = \psi_{0,ij} \eta_{j,t} + \psi_{1,ij} \eta_{j,t-1} + \psi_{2,ij} \eta_{j,t-2} + \psi_{3,ij} \eta_{j,t-3} \ldots \]

More details: see p. 101-103 of Dieppe et al. (2016) - link
**Forecasting with VAR models**

**VAR(p) model:**
\[ y_t = C + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \epsilon_t, \epsilon_t \sim N(0, \Sigma) \]

**Point forecast for horizon \( h \):**
\[ y_{t+h|t} = C + A_1 y_{t+h-1|t} + A_2 y_{t+h-2|t} + \cdots + A_p y_{t+h-p|t} \]

**VMA representation:**
\[ y_t = \Psi_0 \epsilon_t + \Psi_1 \epsilon_{t-1} + \Psi_2 \epsilon_{t-2} + \Psi_3 \epsilon_{t-3} + \cdots \]

**Error of forecast for horizon \( h \) due to future shocks**
\[ y_{t+h} - y_{t+h|t} = \Psi_0 \epsilon_{t+h} + \Psi_1 \epsilon_{t+h-1} + \cdots + \Psi_{h-1} \epsilon_{t+1} \]

**Variance of forecast error for horizon \( h \) due to future shocks**
\[ Var_t(y_{t+h}) = \Psi_0 \Sigma \Psi'_0 + \Psi_1 \Sigma \Psi'_1 + \cdots + \Psi_{h-1} \Sigma \Psi'_{h-1} \]
Exercises

Exercise 3.1.
For the model ($y$ and $y^*$ denote output at home and abroad):

$$
\begin{bmatrix}
y^*_t \\
y_t
\end{bmatrix} = 
\begin{bmatrix}
0.25 \\
0.50
\end{bmatrix} + 
\begin{bmatrix}
0.75 & 0.00 \\
0.25 & 0.50
\end{bmatrix} 
\begin{bmatrix}
y^*_{t-1} \\
y_{t-1}
\end{bmatrix} + 
\begin{bmatrix}
1.00 & 0.00 \\
0.50 & 1.00
\end{bmatrix} 
\begin{bmatrix}
\eta^*_t \\
\eta_t
\end{bmatrix}, \eta_t \sim N(0, I)
$$

- Check if the model is stationary
- Calculate the equilibrium value of $y_t$ and $y^*_t$
- Knowing $y^*_T = 1$ and $y_T = 2$ calculate the forecast for periods $T + 1$ and $T + 2$.
- Write the model in $VMA(\infty)$ form. Calculate the values for the first two lags.
- Calculate FEVD for $y_{T+1}$ and $y_{T+2}$
- Calculate point and 95% interval forecast for $y_{T+1}$

Exercise 3.2.
For a model $VAR(2)$

$$
\begin{bmatrix}
y^*_t \\
y_t
\end{bmatrix} = 
\begin{bmatrix}
0.25 \\
0.50
\end{bmatrix} + 
\begin{bmatrix}
0.50 & 0.00 \\
0.25 & 0.25
\end{bmatrix} 
\begin{bmatrix}
y^*_{t-1} \\
y_{t-1}
\end{bmatrix} + 
\begin{bmatrix}
0.25 & 0.00 \\
0.00 & 0.25
\end{bmatrix} 
\begin{bmatrix}
y^*_{t-1} \\
y_{t-1}
\end{bmatrix} + 
\begin{bmatrix}
1.00 & 0.00 \\
0.50 & 1.00
\end{bmatrix} 
\begin{bmatrix}
\eta^*_t \\
\eta_t
\end{bmatrix}, \eta_t \sim N(0, I)
$$

- build a companion matrix
- calculate [in R] if it is stationary
- compute [in R] $VMA$ representation for the first four lags
Exercise 3.3.
Select an EU country \(y_t\) and euro area \(y_t^*\) and a variable (annual inflation, annual GDP growth rate, unemployment rate or gov. bond 10Y yield) and:
A. Estimate VAR model for \(Y_t = [y_t^* \ y_t]^\prime\) (select lags, check for autocorrelation)
B. Identify monetary policy shock (Cholesky decomposition)
C. Plot IRF
D. Plot FEVD
E. Calculate FEVD from IRF?
F. Calculate historical decomposition
G. Make a forecast for the next two years

Exercise 3.4.
Select an EU country and download data for changes in output \(\Delta y_t\) and the unemployment rate \(u_t\):
A. Estimate VAR model for \(Y_t = [\Delta y_t \ u_t]^\prime\), make BQ long-term structuralization
B. Plot IRF, FEVD, historical decomposition
C. Calculate output gap and compare it to data from AMECO database (link)
Topic 4. Out-of-sample forecast evaluation (backtesting)
About forecasting

The ultimate goal of a positive science is to develop a theory or hypothesis that yields valid and meaningful predictions about phenomena not yet observed. Theory is judged by its predictive power.

A hypothesis can't be tested by its assumptions. What is important is specifying the conditions under which the hypothesis works. What matters is its predictive power, not its conformity to reality.

About forecasting

- Predicting future economic outcomes is helpful in making appropriate plans, making investment decisions, conducting economic policies.

- We make inference about future outcome using available data (for time series: current and past data) and statistical models. We call this process **econometric forecasting**

- Point forecast from model $M$, horizon $h$ and information set $\Omega_T$:

  $$y_{T,h} = y_{T+h|T} = E_T(y_{T+h}|M) = E(y_{T+h}|M, \Omega_T)$$

- Density forecast provides information on all quantiles of the distribution. We focus on the entire distribution (pdf):

  $$p_{t,h}^f(u) = p_{t+h|t}(u)$$
Example of forecast for inflation in Poland

Source: Narodowy Bank Polski, Inflation Report
About forecasting

Types of time series forecasts

- Qualitative / model-based (e.g. from VAR/DSGE model)
- Quantitative / expert based (e.g. survey forecast, SPF)

General characteristics of time series forecasts:

- Forecasting is based on the assumption that the past predicts the future
  
  *Think carefully if the past is related to what you expect about the future*

- Forecasts are always wrong
  
  *However, some models/methods might work better or worse than the other*

- Forecasts are usually more accurate for shorter time periods
  
  *But, economic theories are more informative for longer horizon*
**ex-ante vs ex-post forecast**

- Ex-ante forecast is a true inference about the future. It is for periods in which we don't know the realization.
- Ex-post forecast is to check model reliability. It is for periods in which we know the realization.
Assume we know DGP, i.e. the parameters and the specification of ARMA/VAR. We therefore know the parameters of infinite moving average representation

$$y_t = \mu + \psi_0 \eta_t + \psi_1 \eta_{t-1} + \psi_2 \eta_{t-2} + \psi_3 \eta_{t-3} \ldots \quad \eta_t \sim N(0, I)$$

Forecast from known DGP is called optimum forecast. We cannot obtain more accurate forecast from another model.

Forecast error of optimal forecast is solely due to futures shocks (random error):

$$y_{T+h} - y_{T+h|T} = \psi_0 \eta_{T+h} + \psi_1 \eta_{T+h-1} + \cdots + \psi_{h-1} \eta_{T+1}$$

The resulting variance of forecast is:

$$E \left\{ (y_{T+h} - y_{T+h|T})^2 \right\} = \psi_0 \psi'_0 + \psi_1 \psi'_1 + \cdots + \psi_{h-1} \psi'_{h-1}$$
Ex-ante forecast  
error in estimated ARMA/VAR model

- Assume that we don't know the true DGP but use a model $M$ instead.
- The variance of our forecast is:

$$E\{ (y_{T+h} - y_{T+h|T}^M)^2 \} = E\{ (y_{T+h} - y_{T+h|T})^2 \} +$$
$$+ E\{ (y_{T+h|T} - y_{T+h|T}^M)^2 \} +$$
$$+ 2E\{ (y_{T+h} - y_{T+h|T})(y_{T+h|T} - y_{T+h|T}^M) \}$$

**Component A:** error of "optimum forecast" (see previous slide)

**Component B:** estimation / misspecification error
we want to minimize this value

**Component C:** equals to 0 if we cannot forecast future shock
Let us focus on the estimation / misspecification error and model complexity

\[ E\{ (y_{T+h|T} - y_{T+h|T}^M)^2 \} \]

I. Large / complex models
   - many parameters = large estimation error (high variance)
   - many explanatory variables = good specification (low bias)

II. Small / simple models
   - few parameters = small estimation error (low variance)
   - few explanatory variables = potential misspecification (high bias)

Which effect dominates? We don't know and need to check it
Ex-ante forecast
Illustration of the variance / bias trade-off
Ex-ante forecast
Illustration of the variance / bias trade-off

- Let us assume that the true DGP is AR(1):

\[ y_t = \mu + \rho (y_{t-1} - \mu) + \epsilon_t \]

- We have a sample of 180 monthly observations (15 years) for \( y_t \) and would like to decide on one of the three competing models:

  - RW, Random walk: \( \mu^{RW} = 0 \) and \( \rho^{RW} = 1 \)
  - HL, 5-year half life model: \( \mu^{HL} = \bar{y} \) and \( \rho^{HL} = 0.5^{1/60} \)
  - AR, estimated AR model: \( \mu^{AR} \) and \( \rho^{AR} \) are estimated

- Which model performs best? It depends on the value of \( \rho \)
Ex-ante forecast
Illustration of the variance / bias trade-off

Fig. 3 Theoretical mean squared forecast errors. Notes: Each line represents the ratio of MSFE from a given method to MSFE from the random walk, where values below unity indicate better accuracy of point forecasts. The straight and dotted lines stand for AR and HL, respectively. The forecast horizon is expressed in months.

Source: Ca' Zorzi M., Mućk J., Rubaszek M., 2016. RER forecasting and PPP: This time the Random Walk loses, Open Economies Review
About ex-post forecast

- We usually work with models that performed well in the past
- In ex-post forecast we ask a question: *how accurate forecasts the model would deliver if it was used in the past*
- We evaluate *ex-post* forecasts to be sure about model reliability
- An important issue is the use of "real time data, RTD"

![Diagram showing in-sample simulation, ex-post forecast, ex-ante forecast, estimation period, and time](image-url)
About *ex-post* forecast

- We compare forecast $y_{t+h}^f$ from model $M_i$ to realization $y_{t+h}$ to assess:
  - the absolute quality of forecasts from model $M_i$
    MFE, efficiency/unbiasedness tests, sequential forecasts, PIT
  - the relative quality of forecasts from models $M_i$ and $M_j$
    RMSFE/MSFE/MAFE, log predictive scores

- Various forecasting schemes
  - rolling scheme
  - recursive schemes
  - fixed schemes

- A very important choice relates to the split of the sample into estimation and evaluation subsamples
Forecasting schemes - illustration

Figure 1(a). Recursive estimation scheme

Figure 1(b). Rolling estimation scheme

Source: Barbara Rossi, 2014. Density forecasts in economics and policymaking, CREI WP 37
Rolling forecasting scheme - illustration

Source: Rubaszek and Skrzypczyński (2008, IJF)
Point forecasts accuracy measures

Mean forecasts error for horizon $h$:

$$MFE_h = \frac{1}{T_h} \sum_{t=T_1+1}^{T-h} (y_{t+h} - y_{t,h}^f)$$

Root mean squared forecast error:

$$RMSFE_h = \sqrt{\frac{1}{T_h} \sum_{t=T_1+1}^{T-h} (y_{t+h} - y_{t,h}^f)^2}$$

where $T_h = T - T_1 - h + 1$

Diebold-Mariano test for equal forecast accuracy:

Forecast errors from two competing models

$$e_{1t,h} = y_{t+h} - y_{1t,h}^f \text{ and } e_{2t,h} = y_{t+h} - y_{2t,h}^f$$

The quadratic loss differential

$$d_{t,h} = e_{1t,h}^2 - e_{2t,h}^2$$

The null of equal forecast accuracy (RMSFE)

$$H_0: E(d_{t,h}) = 0$$

Test statistic:

$$DM = \frac{\bar{d}_{t,h}}{\sqrt{S/T_h}} \sim N(0,1)$$

where $S = \sum_{i=-(h-1)}^{h-1} \hat{\gamma}(i)$ is the "long-term" variance
### Point fct. accuracy measures: illustration

#### Mean Forecast Errors (MFEs) of Unconditional Forecasts

<table>
<thead>
<tr>
<th></th>
<th>DSGE</th>
<th>SPF</th>
<th>DSGE-VAR((\lambda))</th>
<th></th>
<th>DSGE-VAR((\infty))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(p = 2)</td>
<td>(p = 4)</td>
<td>(p = 6)</td>
</tr>
<tr>
<td>0</td>
<td>-0.57**</td>
<td>0.38</td>
<td>-0.98***</td>
<td>-0.98***</td>
<td>-0.86***</td>
</tr>
<tr>
<td>1</td>
<td>-0.24</td>
<td>0.17</td>
<td>-1.01***</td>
<td>-0.98***</td>
<td>-0.87**</td>
</tr>
<tr>
<td>2</td>
<td>-0.03</td>
<td>0.12</td>
<td>-1.05***</td>
<td>-0.94**</td>
<td>-0.74*</td>
</tr>
<tr>
<td>3</td>
<td>0.11</td>
<td>0.02</td>
<td>-1.07***</td>
<td>-0.86**</td>
<td>-0.66*</td>
</tr>
<tr>
<td>4</td>
<td>0.07</td>
<td>-0.19</td>
<td>-1.18***</td>
<td>-0.88**</td>
<td>-0.72*</td>
</tr>
</tbody>
</table>

#### Root Mean Squared Forecast Errors (RMSFEs) of Unconditional Forecasts

<table>
<thead>
<tr>
<th></th>
<th>DSGE</th>
<th>SPF</th>
<th>DSGE-VAR((\lambda))</th>
<th></th>
<th>DSGE-VAR((\infty))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(p = 2)</td>
<td>(p = 4)</td>
<td>(p = 6)</td>
</tr>
<tr>
<td>0</td>
<td>1.95</td>
<td>0.98</td>
<td>1.10</td>
<td>1.17**</td>
<td>1.15</td>
</tr>
<tr>
<td>1</td>
<td>1.99</td>
<td>1.07</td>
<td>1.18</td>
<td>1.23**</td>
<td>1.23*</td>
</tr>
<tr>
<td>2</td>
<td>1.83</td>
<td>1.21**</td>
<td>1.23*</td>
<td>1.26*</td>
<td>1.25*</td>
</tr>
<tr>
<td>3</td>
<td>1.89</td>
<td>1.19**</td>
<td>1.23**</td>
<td>1.21*</td>
<td>1.19*</td>
</tr>
<tr>
<td>4</td>
<td>2.10</td>
<td>1.16**</td>
<td>1.21**</td>
<td>1.16**</td>
<td>1.17*</td>
</tr>
</tbody>
</table>

Source: Kolasa, Rubaszek, Skrzypczynski (2012, JMBC)
Point fct. accuracy measures: efficiency

Efficiency / unbiasedness test

A relatively good forecast accuracy does not imply that they are satisfactory in the absolute sense! Absolute performance include ME and efficiency/unbiasedness test. For regression:

$$y_{t+h} = \alpha_0 + \alpha_1 y_{t,h}^f + \varepsilon_{t,h}$$

we test whether $\alpha_0 = 0$ and $\alpha_1 = 1$.

[ the alternative specification is $e_{t,h} = \alpha_0 + \alpha_1 y_{t,h}^f + \varepsilon_{t,h}$ in which we test $\alpha_0 = 0$ and $\alpha_1 = 0$ ]

Source: Kolasa, Rubaszek, Skrzypczynski (2012, JMBC)
Point forecasts accuracy measures

Efficiency / unbiasedness test – graphical illustration

Fig. 2. Actuals and Four-Quarter-Ahead Forecasts.

Point forecasts: sequential forecasts

Fig. 3. Sequential real exchange rate forecasts.

Source: Ca’Zorzi, Kolas, Rubaszek (2017, JIE)
**Density forecasts accuracy: PIT**

PIT – probability Integrat Transform

\[ PIT_{t,h} = \int_{-\infty}^{y_{t+h}} p_{t,h}^f(u) du \]

where \( p_{t,h}^f() \) is the forecast for density distribution.

For a well calibrated model the series \( PIT_{t,h} \) should be drawn from \( IID \ U(0,1) \)

Source: Kolasa, Rubaszek, Skrzypczynski (2012, JMBC)
Density forecasts accuracy: predictive scores

LPS – log predictive score

\[ LPS_{t,h} = \log(p_{t,h}^f(y_{t+h})) \]

where \( p_{t,h}^f() \) is the forecast for density distribution.

We can compare density forecasts from two models with the Amisano and Giacomini (2007) test of equal forecast accuracy:

The loss differential
\[ L_{t,h} = LPS_{1t,h} - LPS_{2t,h} \]

The null of equal forecast accuracy
\[ H_0: E(L_{t,h}) = 0 \]

Test statistic:
\[ GA = \frac{L_{t,h}}{\sqrt{S/T_h}} \rightarrow N(0,1) \]

where \( S \) is the HAC (Newey and West) estimator of the "long-term" variance for \( L_{t,h} \)

Table: Log Predictive Scores (LPS) for DSGE models

<table>
<thead>
<tr>
<th>H</th>
<th>United Kingdom</th>
<th></th>
<th></th>
<th></th>
<th>Canada</th>
<th></th>
<th></th>
<th></th>
<th>Australia</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>-0.15*</td>
<td>0.04*</td>
<td>-0.07*</td>
<td>-0.08*</td>
<td>-0.39*</td>
<td>-0.05</td>
<td>-0.18°</td>
<td>-0.06*</td>
<td>0.01</td>
<td>0.03*</td>
<td>-0.06*</td>
<td>-0.09*</td>
</tr>
<tr>
<td>2</td>
<td>-0.13°</td>
<td>0.05</td>
<td>-0.06</td>
<td>-0.08*</td>
<td>-0.26*</td>
<td>-0.03</td>
<td>-0.10</td>
<td>0.00</td>
<td>0.03</td>
<td>0.04*</td>
<td>-0.06*</td>
<td>-0.12*</td>
</tr>
<tr>
<td>4</td>
<td>-0.14°</td>
<td>0.03</td>
<td>-0.10</td>
<td>-0.12*</td>
<td>-0.07</td>
<td>0.01</td>
<td>-0.11°</td>
<td>0.02</td>
<td>0.09°</td>
<td>0.07*</td>
<td>-0.06*</td>
<td>-0.19*</td>
</tr>
<tr>
<td>6</td>
<td>-0.17°</td>
<td>0.02</td>
<td>-0.17</td>
<td>-0.16°</td>
<td>0.06</td>
<td>0.02</td>
<td>-0.16°</td>
<td>0.02</td>
<td>0.15°</td>
<td>0.11°</td>
<td>-0.07*</td>
<td>-0.24*</td>
</tr>
<tr>
<td>8</td>
<td>-0.22°</td>
<td>0.02</td>
<td>-0.24</td>
<td>-0.20*</td>
<td>0.12°</td>
<td>0.03</td>
<td>-0.21*</td>
<td>0.01</td>
<td>0.22°</td>
<td>0.14°</td>
<td>-0.08*</td>
<td>-0.28*</td>
</tr>
<tr>
<td>12</td>
<td>-0.34°</td>
<td>0.01</td>
<td>-0.37*</td>
<td>-0.29°</td>
<td>0.19°</td>
<td>0.04</td>
<td>-0.30°</td>
<td>-0.01</td>
<td>0.33°</td>
<td>0.20°</td>
<td>-0.12°</td>
<td>-0.31°</td>
</tr>
</tbody>
</table>

Notes: The figures in the table represent the differences of the LPS from a given model in comparison to the NK benchmark so that positive values indicate that forecasts from a given NOEM variant are more accurate than from the benchmark. Asterisks *, *, and ° denote, respectively, the 1%, 5% and 10% significance levels of the two-tailed Amisano and Giacomini (2007) test, where the long-run variance is calculated with the Newey-West method.

Source: Kolasa and Rubaszek (2018, IJF)
Exercise 4.1.

Select an EU country a variable of interest (inflation, unemployment, output growth) and:

A. Calculate recursive point forecasts from RW, ARMA, VAR models over the last 3 years
B. Calculate MFE and RMSFE for the 3 methods
C. Compare the accuracy of forecasts from 3 models to RW with DM test
D. Conduct efficiency test and draw a scatter-plot for forecasts from the VAR
E. Make a plot for sequential forecasts from VAR and BVAR models
F. Discuss the results
BLOCK 1. Presentation
Plan of the presentation
[max 12p to be gained]

Select an EU country and a variable among:
- annual inflation,
- annual GDP growth rate,
- unemployment rate or
- gov. bond 10Y yield:

A. [1p] Describe the variable
   To show: time series plot, ACF, UR test

B. [2p] Estimate the ARMA model.
   To show: information criteria, IRF.

C. [3p] Estimate the VAR model for a vector \([y_t^*, y_t]\), where \(y_t^*\) is the value of the variable for the euro area, and perform Cholesky structuralization.
   To show: IRF, FEVD, historical decomposition

D. [2p] Compare the accuracy of forecasts from RW, ARMA and VAR models
   To show: MFE, RMSFE, DM test, sequential forecasts graph

E. [2p] Plot forecast from ARMA, VAR and for the next two years and from European Comission (to be found on the webpage)
   To show: A table with forecasts, a graph with three forecasts

About presentation:
- I attribute up to 2p for the quality of the presentation
- presentation should take up to 8 minutes [longer presentations: penalty -1p]