Testing methodology

- It often the case that we try to determine the form of the model on the basis of data

- The simplest case: we try to determine the set of explanatory variables in the model

- Testing for significance of the set of K variables we can test:

1. Joint hypothesis $H_0 : \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_K = 0$, at significance level $\boldsymbol{\alpha}$.

2. K simple hypotheses $H_1 : \boldsymbol{\beta}_1 = 0; \ldots; H_K : \boldsymbol{\beta}_K = 0$, at significance level of $\boldsymbol{\alpha}$ for each of them

- For the second procedure we reject the null if one of the hypotheses $H_1, \ldots, H_K$ is rejected

- This two testing procedures are not equivalent

- Assume that test statistics are independent

- In the second procedure true significance level is equal to:

$$\boldsymbol{\alpha}^* = 1 - (1 - \boldsymbol{\alpha})^K$$

- Notice that for $\lim_{K \longrightarrow \infty} \boldsymbol{\alpha}^* = 1$!

Conclusion 1. *we should always test the joint hypothesis rather than separately test the simple hypotheses. Otherwise we should make adjustments to significance level.*

- Difference between nominal significance level α and true significance level α^* is called Lovell bias.

General to specific modelling

- Objective: to find the correct specification of the model on the basis of data

- Searching for correct specification of the model it is often possible to formulate the series of nested models and hypotheses

Example 2. *Determining the set of explanatory variables:*

- - $H_0^0 : \boldsymbol{\beta}_i \neq 0$ for $i = 1, \ldots, K$
  - $H_0^1 : \boldsymbol{\beta}_1 = 0$
  - $H_0^2 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = 0$
  - $\vdots$
  - $H_0^K : \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_K = 0$

- H_0^K is nested in H_0^1 and H_0^2, if it imposes some restricts over the ones imposed H_0^1 and H_0^2.

- Each time we test H_0^i under alternative H_0^0 (usually with F test)

- We stop the imposing restrictions when H_0^i is rejected, we choose model given by H_0^{i-1} as correct the model

Information criteria

- Information criteria are used to compare the models and choose the best one
- Fit of the model cannot be the base of the choice between models because for larger number of parameters (variables), fit is always better
- Conventionally the information criteria are defined in such a way that the best model has the lowest information criterion is the best
- Information criteria are defined in such a way that they take into account the fact that better fit can always be achieved with more parameters. They only improve if the improvement in fit is ”significant”.

- The most often used are **A**kaike **I**nformation **C**riterion (AIC) and **B**ayes **I**nformation **C**riterion (BIC)

$$BIC = -\frac{2\ell\left(\widehat{\boldsymbol{\theta}}\right)}{n} + \frac{K\log(n)}{n}$$
$$AIC = -\frac{2\ell\left(\widehat{\boldsymbol{\theta}}\right)}{n} + \frac{2K}{n}$$

Dynamic models

- Objective: to describe the behavior of the economic system in time
- Following dynamic characteristics are of the interest to economists
  - existence and parameters of the long run equilibrium
  - how fast the system adjust to long run equilibrium
  - what is the time pattern of reactions of the economic system to exogenous shocks (e.g. policy changes)
  - seasonality
- Additional reason to investigate the dynamic features of the economic system is the analysis of causality and forecasting

- Forecasting is only possible if we are able to identify variables (causes), whose changes with some time lag influence the other variables
- In dynamic models we always try to eliminate autocorrelation because it can bias the estimates of the parameters and also is a signal that we failed to explain the dynamics of the dependent variable

Distributed lag model (DL)

- Model in which only the DL (**D**istributed **L**ags) part is present

- It is simple to analyze because $\boldsymbol{x}_t$ and lagged $\boldsymbol{x}_t$ can be assumed to be exogenous

- Model can satisfy the assumptions of Classical Regression Model

$$y_t = \boldsymbol{\mu} + \boldsymbol{\beta}_0 \boldsymbol{x}_t + \ldots + \boldsymbol{\beta}_p \boldsymbol{x}_{t-p} + \boldsymbol{\varepsilon}_t$$

- Coefficients of explanatory variables describe the reaction of the y on the changes of x in time t but also in earlier periods

- When interpret this coefficients it is important to clarify whether we mean:
  - short run reaction of y_t to the unit change of x_t (impact multiplier)
  - cumulated reaction of y_t to the unit change of $x_t, \ldots, x_{t-\boldsymbol{\tau}}$ (cumulated multiplier)
  - the long run reaction of y_t to the permanent unit change of x_t (long-run multiplier)

- For DL models:
  - Change of $\boldsymbol{x}_t$ by $\Delta \boldsymbol{x}_t$ causes the change of y_t equal to:

$$\begin{aligned} \mathrm{E}\left(y_t + \Delta y_t\right) &= \boldsymbol{\mu} + \boldsymbol{\beta}_0 \left(\boldsymbol{x}_t + \Delta \boldsymbol{x}_t\right) + \ldots + \boldsymbol{\beta}_p \boldsymbol{x}_{t-p} \\ &= \mathrm{E}\left(y_t\right) + \boldsymbol{\beta}_0 \Delta \boldsymbol{x}_t \end{aligned}$$

Impact multiplier is then equal to:

$$\frac{\mathrm{E}\left(\Delta y_t\right)}{\Delta \boldsymbol{x}_t} = \boldsymbol{\beta}_0$$

– Change of $\boldsymbol{x}_t$ by $\Delta \boldsymbol{x}_t$ which happened $\boldsymbol{\tau}$ periods before t and influenced $\boldsymbol{x}_t$ afterwards causes the change of y_t equal to:

$$\begin{aligned}
\mathrm{E}\left(y_t + \Delta y_t\right) &= \boldsymbol{\mu} + \boldsymbol{\beta}_0\left(\boldsymbol{x}_t + \Delta \boldsymbol{x}\right) + \ldots + \boldsymbol{\beta}_{\boldsymbol{\tau}}\left(\boldsymbol{x}_{t-\boldsymbol{\tau}} + \Delta \boldsymbol{x}\right) \\
&\quad + \boldsymbol{\beta}_{\boldsymbol{\tau}+1}\boldsymbol{x}_{t-\boldsymbol{\tau}+1} + \ldots + \boldsymbol{\beta}_p \boldsymbol{x}_{t-p} \\
&= \mathrm{E}\left(y_t\right) + \left(\sum_{i=0}^{\boldsymbol{\tau}} \boldsymbol{\beta}_i\right)\Delta \boldsymbol{x}
\end{aligned}$$

Cumulated multiplier is then equal to:

$$\boldsymbol{\beta}_{\boldsymbol{\tau}} = \frac{\mathrm{E}\left(\Delta y_{t+\boldsymbol{\tau}}\right)}{\Delta \boldsymbol{x}} = \sum_{i=0}^{\boldsymbol{\tau}} \boldsymbol{\beta}_i$$

– Long run influence of the permanent change of $\boldsymbol{x}$ by $\Delta \boldsymbol{x}$ is measured with long run multiplier. It is equal to cumulated multiplier for $\boldsymbol{\tau} \to \infty$

$$\frac{\mathrm{E}\left(\Delta y\right)}{\Delta \boldsymbol{x}} = \boldsymbol{\beta} = \sum_{i=0}^{\infty} \boldsymbol{\beta}_i$$

- Speed of reaction of the dependent variable to changes of independent variables can be measured with mean lag: $\overline{w} = \sum_{i=1}^{\infty} i \frac{\boldsymbol{\beta}_i}{\boldsymbol{\beta}}$

Exercise 3. *Relationship between unemployment according to BAEL (ILO definition) and supply of money in nominal terms ($m3p$) - Polish quarterly data*

Choice of the lag length - general to specific method: it was assumed that the maximum sensible number of lags was 6

```
------------------------------------------------------------------------------
bael         |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
m3p          |
          -- |  -.2284744   .1149244    -1.99   0.056    -.4635211    .0065723
          L1 |  -.2218003   .1192405    -1.86   0.073    -.4656745    .0220739
          L2 |  -.2727794   .1166159    -2.34   0.026    -.5112856   -.0342732
          L3 |  -.2606365   .0965898    -2.70   0.011    -.4581849   -.0630882
          L4 |  -.1721529   .1082111    -1.59   0.122    -.3934693    .0491636
          L5 |  -.0142759   .1108159    -0.13   0.898    -.2409199    .2123681
          L6 |  -.0176795   .1072485    -0.16   0.870    -.2370272    .2016683
_cons        |   20.51169   .5000556    41.02   0.000     19.48897    21.53442
------------------------------------------------------------------------------
```

- $\beta_6 = 0$: F(1, 29) = 0.03 [0.8702]

  AIC: 3.945

  BIC: 25.231

- $\beta_6 = \beta_5 = 0$: F(2, 29) = 0.02[0.9825]

  AIC: 3.904

  BIC: 21.588

- $\beta_6 = \beta_5 = \beta_4 = 0$: F(3, 29) = 0.90 [0.4513]

  AIC: 3.951

  BIC: 21.188

- $\beta_6 = \beta_5 = \beta_4 = \beta_3 = 0$: F(4, 29) = 2.46[0.0673]

AIC: 4.088

BIC: 24.396

- Information criterion AIC suggests 4 lags, BIC suggests 3 lags
- Testing from general to specific at $\boldsymbol{\alpha} = 10\%$ - suggests 3 lags
- We choose 3 lags

```
      Source |       SS       df       MS              Number of obs =      40
-------------+------------------------------           F(  4,    35) =   30.27
       Model |  376.021652     4  94.005413            Prob > F      =  0.0000
    Residual |  108.707887    35  3.10593963           R-squared     =  0.7757
-------------+------------------------------           Adj R-squared =  0.7501
       Total |  484.729539    39  12.4289625           Root MSE      =  1.7624

------------------------------------------------------------------------------
bael         |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
m3p          |
         --  |  -.2819283   .0991632    -2.84   0.007    -.4832404   -.0806162
         L1  |  -.2017262   .0938015    -2.15   0.038    -.3921533   -.0112991
         L2  |  -.2861484   .0944252    -3.03   0.005    -.4778417    -.094455
         L3  |  -.2836885     .09924    -2.86   0.007    -.4851565   -.0822205
_cons        |   20.05402   .5144917    38.98   0.000     19.00954    21.09849
------------------------------------------------------------------------------
```

- According to IS/LM monetary expansion should reduce unemployment
- Impact multiplier of the change of supply of money (change of

unemployment in reaction of 1% increase in money supply):

$$\boldsymbol{\beta}_0 = -.2819283$$

- Long run multiplier of the change of supply of money (change of unemployment if the rate of money expansion is permanently increased by 1%)

$$\boldsymbol{\beta} = -.2819283 - .2017262 - .2861484 - .2836885 = -1.0535$$

- Long run multiplier is 4 times larger than impact multiplier!

- Mean lag:

$$\overline{w} = \frac{1}{1.0535}(.2017262 + 2 \times .2861484 + 3 \times .2836885) = 1.5426$$

- Because model is based in quarterly data $\overline{w}$ suggests that mean lag of the reaction of unemployment is equal to about 1.5 quarters.

- However: we failed to account for all the dynamics of the reaction of unemployment to money expansion - this can be inferred from the result of the autocorrelation test:

```
Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |         20.953               1                   0.0000
---------------------------------------------------------------------------
                        H0: no serial correlation
```