Elements of Econometrics.
Lecture 24.
Revision Time Series 1.

FCS, 2022-2023

Basic Regression Analysis with Time Series Data

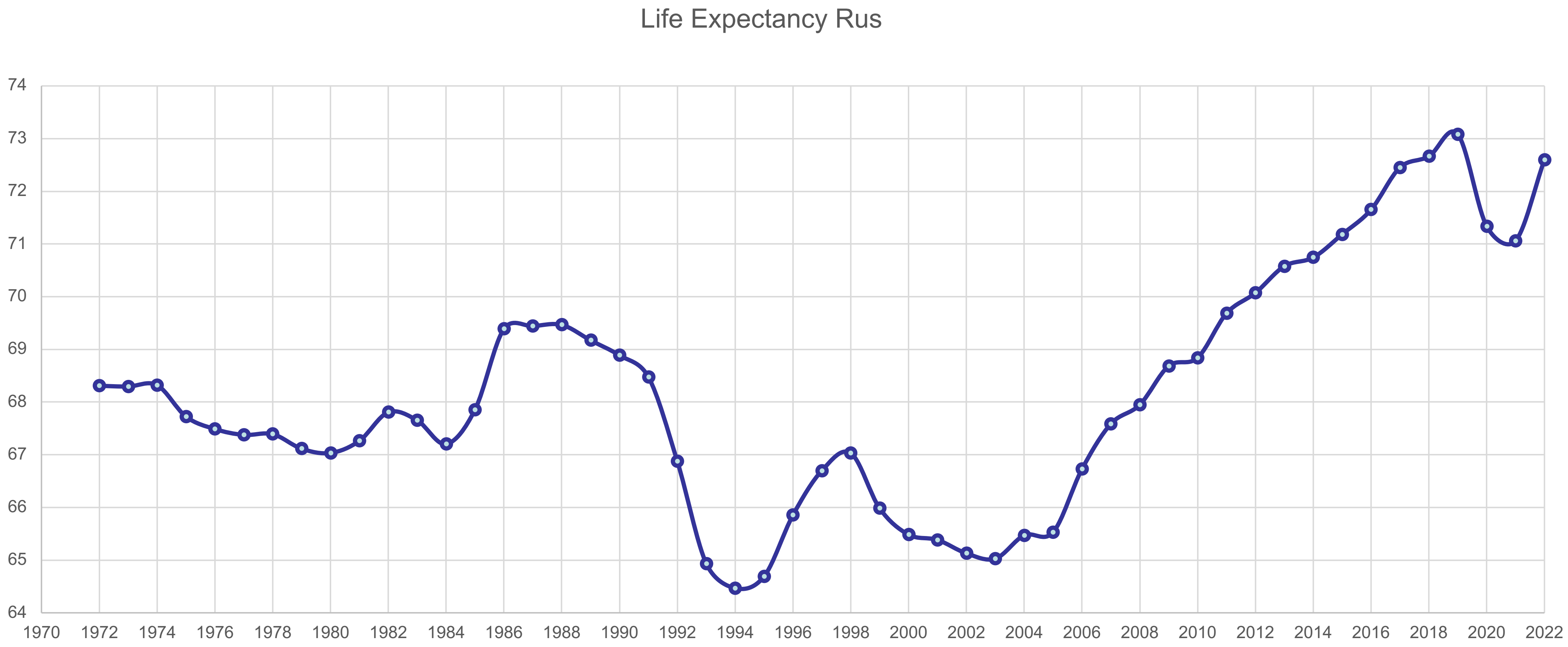
- **The nature of time series data**
- Temporal ordering of observations; may not be arbitrarily reordered
- Typical features: serial correlation/nonindependence of observations

- How should we think about the randomness in time series data?
  - The outcome of economic variables (e.g. GDP, Inflation Rates) is uncertain; they should therefore be modeled as random variables.
  - Time series are sequences of r.v. (= stochastic processes)
  - Randomness does not come from sampling from a population.
  - "Sample" = the one realized path of the time series out of the many possible paths the stochastic process could have taken.

Time Series Data Specifics Essential for Regression Models

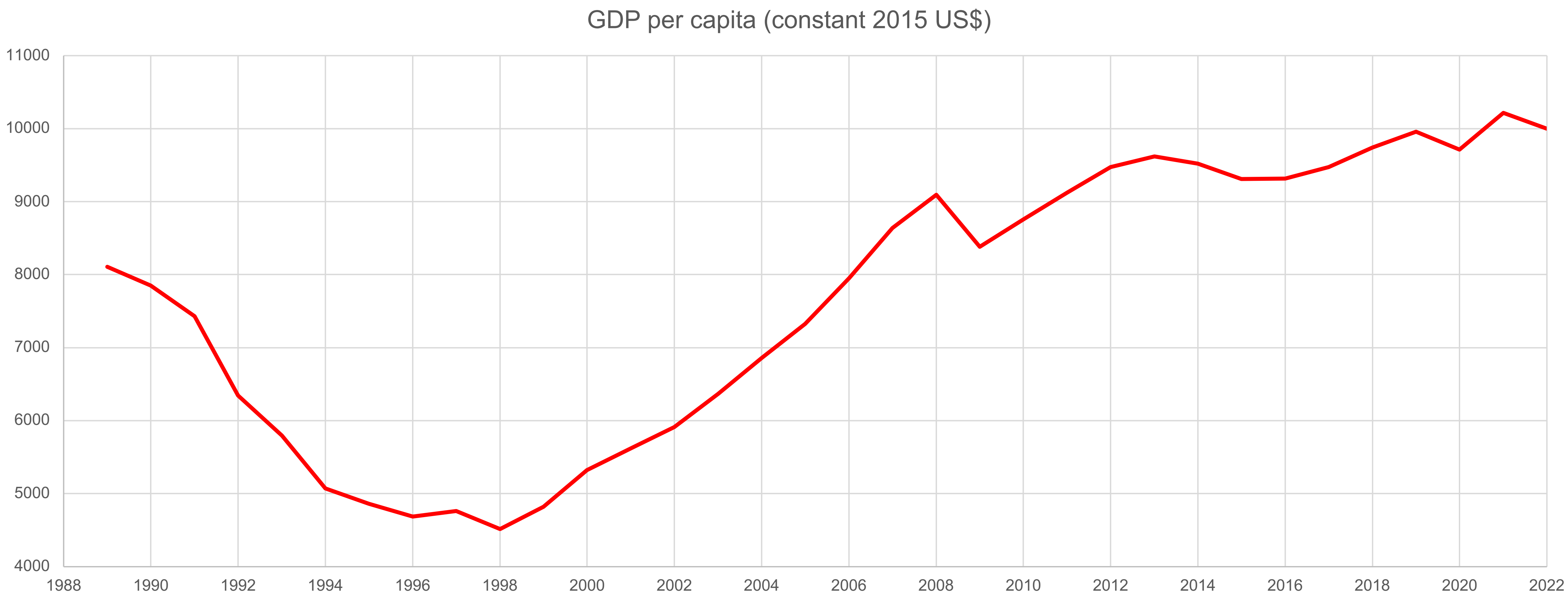
- Assumption B.2 is irrelevant (the observations do not look as being taken randomly from fixed populations): B.2 to be replaced by another assumption
- There may be regularities in the time series and in their relationships: trends, seasonalities, autocorrelations in variables and disturbance terms, lags (fixed or distributed); to be identified and dealt with
- Some regularities in the data (nonstationarity) may lead to estimation of spurious regressions: the data/model has to be transformed to provide desirable estimators' properties

Example: Life Expectancy in Russia, 1972-2022, years.



What are regularities and Irregularities here?

Example: Real GDP Per Capita in Russia, 1989-2022



What are regularities and Irregularities here?

ASSUMPTIONS FOR MODEL C: REGRESSIONS WITH TIME SERIES DATA

ASSUMPTIONS FOR MODEL C

C.1 The model is linear in parameters and correctly specified

$$Y = \beta_1 + \beta_2 X_2 + \ldots + \beta_k X_k + u$$

C.2 The time series for the regressors are (at most) weakly persistent

C.3 There does not exist an exact linear relationship among the regressors

C.4 The disturbance term has zero expectation

C.5 The disturbance term is homoscedastic

C.6 The values of the disturbance term have independent distributions:
u_t is distributed independently of $u_{t'}$ for $t' \neq t$

C.7 The disturbance term is distributed independently of the regressors:
u_t is distributed independently of $X_{jt'}$ for all t' (including t) and j

C.8 The disturbance term has a normal distribution

Assumptions C.1, C.3, C.4, C.5, C.6, C.7 and C.8, and the consequences of their violations are the same as those for Model B. Weak persistency (or weak non-stationarity), C.2: see the next slide.

Basic Regression Analysis with Time Series Data

- **Examples of time series regression models**
- Static models
  - In static time series models, the current value of one variable is modeled as the result of the current values of explanatory variables

- Examples for static models

$$inf_t = \beta_0 + \beta_1 unem_t + u_t$$

There is a contemporaneous relationship between unemployment and inflation (= Phillips curve).

$$mrdrte_t = \beta_0 + \beta_1 convrte_t + \beta_2 unem_t + \beta_3 yngmle_t + u_t$$

The current murder rate is determined by the current conviction rate, unemployment rate, and the fraction of young males in the population.

Basic Regression Analysis with Time Series Data

- **Finite distributed lag models**
  - In finite distributed lag models, the explanatory variables are allowed to influence the dependent variable with a time lag.
- Example for a finite distributed lag model
  - The fertility rate may depend on the tax value of a child, but for biological and behavioral reasons, the effect may have a lag.

$$gfr_t = \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + u_t$$

Children born per
1,000 women in year t

Tax exemption
in year t

Tax exemption
in year t - 1

Tax exemption
in year t - 2

Basic Regression Analysis with Time Series Data

- **Interpretation of the effects in finite distributed lag models**

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \ldots + \delta_q z_{t-q} + u_t$$

- Effect of a past shock on the current value of the dep. variable

$$\frac{\Delta y_t}{\Delta z_{t-s}} = \delta_s$$

$$\frac{\Delta y_t}{\Delta z_{t-q}} + \ldots + \frac{\Delta y_t}{\Delta z_t} = \delta_1 + \ldots + \delta_q$$

<u>Effect of a transitory shock:</u>
If there is a one time shock in a past period, the dep. variable will change temporarily by the amount indicated by the coefficient of the corresponding lag.

<u>Effect of permanent shock:</u>
If there is a permanent shock in a past period, i.e. the explanatory variable permanently increases by one unit, the effect on the dep. variable will be the cumulated effect of all relevant lags. This is a long-run effect on the dependent variable.

Stationary Stochastic Processes

Stationarity (strong stationarity) of a stochastic process X_t is observed if the joint distribution of $X_{t1},X_{t2},\ldots,X_{tm}$ is identical to the joint distribution of $X_{t1+t},X_{t2+t},\ldots,X_{tm+t}$ for any $m,t,t_1,\ldots,t_m$.

A stochastic process is weakly stationary (or covariance stationary) if $E(X_t)$ is constant, $Var(X_t)$ is constant, and for any $t,s \geq 1$, $Cov(X_t, X_{t+s})$ depends only on s and not on t

If for a weakly stationary process $Cov(X_t,X_{t+s}) \rightarrow 0$ as $s \rightarrow \infty$, the process is called weakly dependent (or weakly persistent)

ASSUMPTIONS FOR MODEL C

C.7 The disturbance term is distributed independently of the regressors

u_t is distributed independently of $X_{jt'}$ for all t' (including t) and j

(1) The disturbance term in any observation is distributed independently of the values of the regressors in the same observation, and

(2) The disturbance term in any observation is distributed independently of the values of the regressors in the other observations.

Assumption C.7, like its counterpart Assumption B.7, is essential for both the unbiasedness and the consistency of OLS estimators. Both parts are required for unbiasedness. However only the first part is required for consistency (as a necessary, but not sufficient, condition).

UNBIASEDNESS

$$b_2^{\mathrm{OLS}} = \beta_2 + \frac{\sum (X_i - \bar{X}) u_i}{\sum (X_i - \bar{X})^2}$$

$$b_2^{\mathrm{OLS}} = \beta_2 + \sum a_i u_i \qquad a_i = \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2}$$

$$\mathrm{E}\left(b_2^{\mathrm{OLS}}\right) = \beta_2 + \sum \mathrm{E}(a_i u_i)$$

$$= \beta_2 + \sum \mathrm{E}(a_i)\,\mathrm{E}(u_i)$$

$$= \beta_2 + \sum \mathrm{E}(a_i) \times 0 = \beta_2$$

It is required that u_i distributed independently of a_i. a_i is a function of *all* of the X values in the sample, not just X_i. So Part (1) of Assumption C.7, that u_i is distributed independently of X_i, is not enough.

ASSUMPTION C.7

CROSS-SECTIONAL DATA:
We assume u_j and X_i are independent ($i \neq j$).
The main issue is whether u_i is independent of X_i.

TIME SERIES DATA:

$$\text{If } Y_t = \beta_1 + \beta_2 Y_{t-1} + u_t$$

$$Y_{t+1} = \beta_1 + \beta_2 Y_t + u_{t+1}$$

The disturbance term u_t is automatically correlated with the explanatory variable Y_t in the next observation.

$$b_2^{\text{OLS}} = \beta_2 + \frac{\sum (X_i - \bar{X}) u_i}{\sum (X_i - \bar{X})^2}$$

$$b_2^{\text{OLS}} = \beta_2 + \sum a_i u_i$$

$$a_i = \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2}$$

$$E\left(b_2^{\text{OLS}}\right) = \beta_2 + \sum E(a_i u_i)$$

$$= \beta_2 + \sum E(a_i) E(u_i)$$

$$= \beta_2 + \sum E(a_i) \times 0 = \beta_2$$

Let $Y_{t\text{-}1}$ be a regressor for Y_t. As a consequence u_i is not independent of a_i and so we cannot write $E(a_i u_i) = E(a_i)E(u_i)$. It follows that the OLS slope coefficient will in general be biased.

ASSUMPTION C.7: CONSISTENCY

$$\boldsymbol{Y_t = \beta_1 + \beta_2 Y_{t-1} + u_t}$$

$$b_2^{\mathrm{OLS}} = \frac{\sum (Y_{t-1} - \bar{Y}_{t-1})(Y_t - \bar{Y})}{\sum (Y_{t-1} - \bar{Y}_{t-1})^2} = \beta_2 + \frac{\sum (Y_{t-1} - \bar{Y}_{t-1})(u_t - \bar{u})}{\sum (Y_{t-1} - \bar{Y}_{t-1})^2}$$

$$\operatorname{plim}\left(\frac{\sum (Y_{t-1} - \bar{Y}_{t-1})(u_t - \bar{u})}{\sum (Y_{t-1} - \bar{Y}_{t-1})^2}\right) = \operatorname{plim}\left(\frac{\frac{1}{n}\sum (Y_{t-1} - \bar{Y}_{t-1})(u_t - \bar{u})}{\frac{1}{n}\sum (Y_{t-1} - \bar{Y}_{t-1})^2}\right)$$

$$= \frac{\operatorname{plim} \frac{1}{n}\sum (Y_{t-1} - \bar{Y}_{t-1})(u_t - \bar{u})}{\operatorname{plim} \frac{1}{n}\sum (Y_{t-1} - \bar{Y}_{t-1})^2} = \frac{\sigma_{Y_{t-1}, u_t}}{\sigma^2_{Y_{t-1}}} = \frac{0}{\sigma^2_{Y_{t-1}}} = 0$$

If Part (1) of Assumption C.7 is valid, the covariance between u_t and Y_{t-1} is zero. We suppose that Part(1) is valid because Y_{t-1} is determined before u_t is generated.

Basic Regression Analysis with Time Series Data

- **Example: Static Phillips curve**

$$\widehat{inf}_t = \underset{(1.72)}{1.42} + \underset{(.289)}{.468}\ unem_t$$

Contrary to theory, the estimated Phillips Curve does not suggest a tradeoff between inflation and unemployment

$$n = 49, R^2 = .053, \bar{R}^2 = .033$$

- **Discussion of the Model assumptions**

  C.1 (Wooldridge -TS.1): The error term contains factors such as monetary shocks, income/demand shocks, oil price shocks, supply shocks, or exchange rate shocks.

  C.3 (Wooldridge - TS.2): A linear relationship might be restrictive, but it should be a good approximation. Perfect collinearity is not a problem as long as unemployment varies over time.

Basic Regression Analysis with Time Series Data

- **Discussion of the Model assumptions**

C.4 TS.3: $E(u_t|unem_1,\ldots,unem_n) = 0 \leftarrow$ Easily violated

$unem_{t-1}\uparrow\ \rightarrow\ u_t\downarrow\ \leftarrow$ For example, past unemployment shocks may lead to future demand shocks which may dampen inflation

$u_{t-1}\uparrow\ \rightarrow\ unem_t\uparrow \leftarrow$ For example, an oil price shock means more inflation and may lead to future increases in unemployment

C.5 TS.4: $Var(u_t|unem_1,\ldots,unem_n) = \sigma^2 \leftarrow$ Assumption is violated if monetary policy is more "nervous" in times of high unemployment

C.6 TS.5: $Corr(u_t,u_s|unem_1,\ldots,unem_n) = 0 \leftarrow$ Assumption is violated if exchange rate influences persist over time (they cannot be explained by unemployment)

C.8 TS.6: $u_t \sim \text{Normal}(0,\sigma^2) \leftarrow$ Questionable

MODELS WITH A LAGGED DEPENDENT VARIABLE

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 Y_{t-1} + u_t \quad - \quad ADL(1,0)$$

$$\bar{Y} = \beta_1 + \beta_2 \bar{X} + \beta_3 \bar{Y}$$

$$\bar{Y} = \frac{\beta_1}{1-\beta_3} + \frac{\beta_2}{1-\beta_3} \bar{X}$$

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_{t-1} + \ldots + \beta_{q+2} X_{t-q} +$$
$$+ \alpha_3 Y_{t-1} + \alpha_4 Y_{t-2} + \ldots + \alpha_{p+2} Y_{t-p} + u_t \quad - \quad ADL(p,q)$$

One of the possible ways of including dynamics in a model is an autoregressive distributed lag model, ADL(*p*, *q*). Here p is the maximum lag in *Y*, *q* is the maximum lag in *X*.

DISTRIBUTED LAG MODELS

Distributed lag model: $$Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_{t-1} + \ldots + u_t$$

The following problems arise when estimating the β_i parameters:

- Multicollinearity. Lagged values of an explanatory variable can be strongly correlated with each other.

- The observations' set being reduced when the number of lags grows.

- \- A large number of parameters is estimated and, therefore, the number of degrees of freedom is reduced.

Solution: to make an assumption concerning the distribution of coefficients with small number of parameters (equivalent to imposing restrictions):

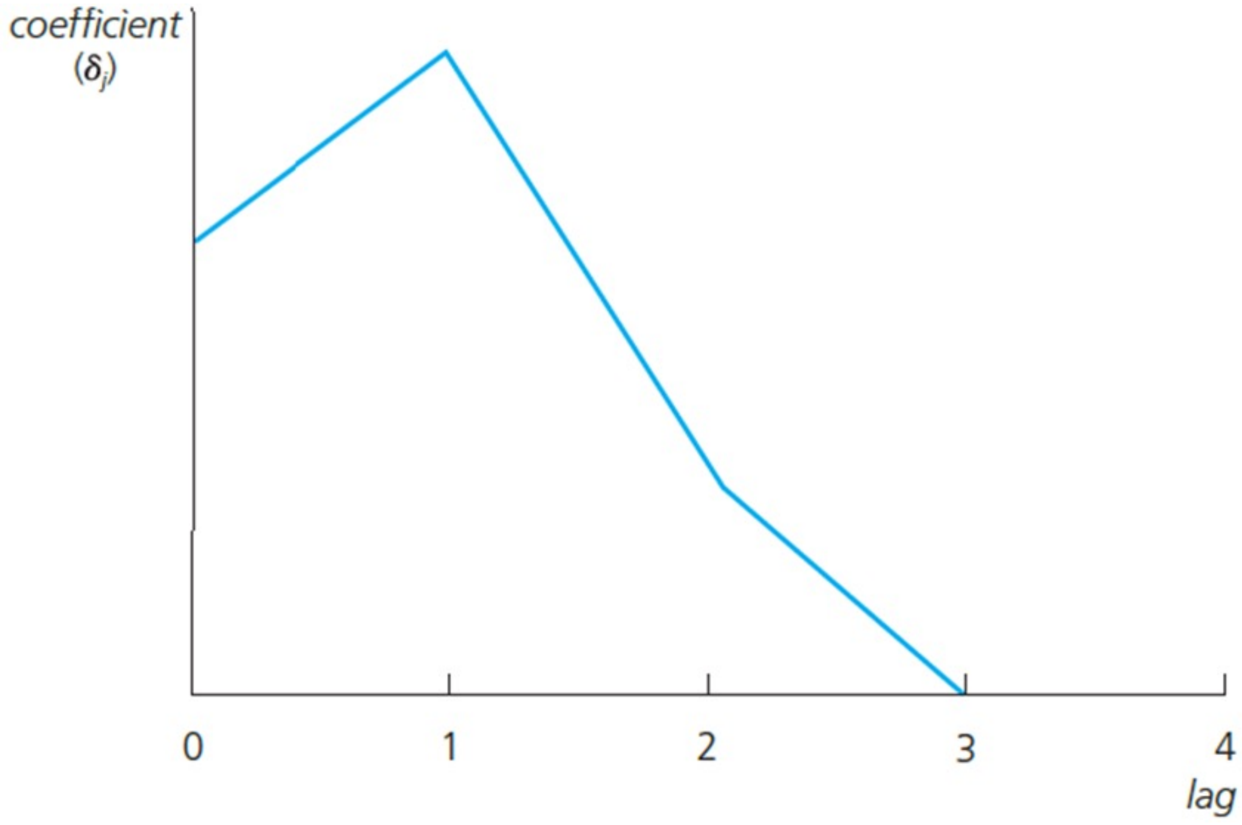
1) Geometric distribution (Koyck distribution) : $$\beta_3 = \beta_2 \cdot \rho; \quad \beta_4 = \beta_2 \cdot \rho^2;$$
where $0 < \rho < 1$

2) Polynomial distribution:
where *s* are integers. $$\beta_s = \tilde{\beta}_0 + \tilde{\beta}_1 \cdot s + \tilde{\beta}_2 \cdot s^2$$

It is also possible to impose on the parameters restrictions of uniform, linear, "triangular" and other distributions.

Basic Regression Analysis with Time Series Data

- **Graphical illustration of lagged effects**



- The effect is biggest after a lag of one period. After that, the effect vanishes (if the initial shock was transitory).

- The long run effect of a permanent shock is the cumulated effect of all relevant lagged effects. It does not vanish (if the initial shock is a permanent one).

GEOMETRICALLY DISTRIBUTED LAG

Geometrically Distributed lag (Koyck model):

$$Y_t = \beta_1 + \beta_2 X_t + \beta_2 \rho X_{t-1} + \beta_2 \rho^2 X_{t-2} + \ldots + u_t$$

Koyck transformation:

$$Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + \beta_2 \rho X_{t-2} + \beta_2 \rho^2 X_{t-3} + \ldots + u_{t-1}$$

$$\Rightarrow Y_t - \rho \cdot Y_{t-1} = \beta_1 (1 - \rho) + \beta_2 \cdot X_t + u_t - \rho \cdot u_{t-1}$$

$$Y_t = \beta_1 (1 - \rho) + \rho \cdot Y_{t-1} + \beta_2 \cdot X_t + v_t$$

If the model is estimated in this form, then the estimates are biased and inconsistent since Y_{t-1} is related with $v_t = u_t - \rho \cdot u_{t-1}$

So the model to be estimated as non-linear regression; lag number increasing stops when the estimates do not change.

MODELS OF DYNAMIC PROCESSES:

PARTIAL ADJUSTMENT MODEL

$$Y_t^* = \beta_1 + \beta_2 X_t + u_t$$

$$\boldsymbol{Y_t - Y_{t-1} = \lambda(Y_t^* - Y_{t-1})}$$

$$\boldsymbol{Y_t = \lambda Y_t^* + (1-\lambda)Y_{t-1}}$$

$$\begin{aligned} Y_t &= \lambda\left(\beta_1 + \beta_2 X_t + u_t\right) + \left(1-\lambda\right)Y_{t-1} \\ &= \beta_1\lambda + \beta_2\lambda X_t + \left(1-\lambda\right)Y_{t-1} + \lambda u_t \\ &= \alpha_1 + \alpha_2 X_t + \alpha_3 Y_{t-1} + \lambda u_t \end{aligned}$$

where $\alpha_1 = \beta_1\lambda,\ \alpha_2 = \beta_2\lambda,\ \alpha_3 = \left(1-\lambda\right).$

Let *Y be target or desired (unobserved) value of *Y.* Parameter *λ* is called the speed of adjustment (complete adjustment if *λ*=1). Substituting for Y_t^* we obtain a specification with observable variables of the ADL(1,0) form.**

TIME SERIES MODELS:

STATIC MODELS AND MODELS WITH LAGS

$$LGHOUS_t{*} = \beta_1 + \beta_2 LGDPI_t + \beta_3 LGPRHOUS_t + u_t$$

$$LGHOUS_t - LGHOUS_{t-1} = \lambda(LGHOUS_t^* - LGHOUS_{t-1})$$

$$LGHOUS_t = \lambda LGHOUS_t^* + (1-\lambda)LGHOUS_{t-1}$$

$$\begin{aligned} LGHOUS_t &= \lambda\left(\beta_1 + \beta_2 LGDPI_t + \beta_3 LGPRHOUS_t + u_t\right) + \left(1-\lambda\right)LGHOUS_{t-1} \\ &= \beta_1\lambda + \beta_2\lambda LGDPI_t + \beta_3\lambda LGPRHOUS_t + \left(1-\lambda\right)LGHOUS_{t-1} + \lambda u_t \\ &= \alpha_1 + \alpha_2 LGDPI_t + \alpha_3 LGPRHOUS_t + \alpha_4 LGHOUS_{t-1} + \lambda u_t \end{aligned}$$

short-run elasticities: $\alpha_2; \quad \alpha_3$

long-run elasticities: $\beta_2 = \dfrac{\alpha_2}{1-\alpha_4} \qquad \beta_3 = \dfrac{\alpha_3}{1-\alpha_4}$

Let *LGHOUS be the desired (unobserved) demand for housing, while the actual demand LGHOUS is determined by the Partial Adjustment process. *LGDPI* is the logarithm of disposable personal income, and *LGPRHOUS* is the logarithm of the relative price index for housing services.**

ADAPTIVE EXPECTATIONS

$$Y_t = \beta_1 + \beta_2 X^e_{t+1} + u_t$$

$$X^e_{t+1} - X^e_t = \lambda\left(X_t - X^e_t\right)$$

$$X^e_{t+1} = \lambda X_t + \left(1-\lambda\right) X^e_t$$

$$Y_t = \beta_1 + \beta_2 \lambda X_t + \beta_2 \left(1-\lambda\right) X^e_t + u_t$$

$$X^e_t = \lambda X_{t-1} + \left(1-\lambda\right) X^e_{t-1}$$

$$Y_t = \beta_1 + \beta_2 \lambda X_t + \beta_2 \lambda \left(1-\lambda\right) X_{t-1} + \beta_2 \left(1-\lambda\right)^2 X^e_{t-1} + u_t$$

$$Y_t = \beta_1 + \beta_2 \lambda X_t + \beta_2 \lambda \left(1-\lambda\right) X_{t-1} + \beta_2 \lambda \left(1-\lambda\right)^2 X_{t-2} + \ldots$$
$$+ \beta_2 \lambda \left(1-\lambda\right)^{s-1} X_{t-s+1} + \beta_2 \left(1-\lambda\right)^s X^e_{t-s+1} + u_t$$

X^e_{t+1} is the expected value of X (unobservable explanatory variable). We suppose that expectations are changing proportionally to the discrepancy between X^e_t, and the actual value X_t.

We assume $0 < \lambda \leq 1$, hence $(1 - \lambda)^s$ tends to zero as s grows, and the term with unobserved variable can be neglected.

THE ERROR CORRECTION MODEL

$$Y_t^* = \beta_1 + \beta_2 X_t + u_t$$

$$\begin{aligned}\Delta Y_t &= \lambda\left(Y_t^* - Y_{t-1}\right) + \delta \Delta X_t \\ &= \lambda\left(\beta_1 + \beta_2 X_t - Y_{t-1}\right) + \delta\left(X_t - X_{t-1}\right) + \lambda u_t \\ &= \lambda\beta_1 + (\lambda\beta_2 + \delta) X_t - \delta X_{t-1} - \lambda Y_{t-1} + \lambda u_t\end{aligned}$$

$$Y_t = \alpha_1 + \alpha_2 X_t + \alpha_3 Y_{t-1} + \alpha_4 X_{t-1} + \lambda u_t$$

$$\alpha_1 = \lambda\beta_1 \qquad \alpha_2 = \lambda\beta_2 + \delta$$

$$\alpha_3 = 1 - \lambda \qquad \alpha_4 = -\delta$$

Y* is a desirable (appropriate) unobserved value of Y. In the short run, $\Delta Y_t = Y_t - Y_{t-1}$, is determined by two components: closing the discrepancy between its "appropriate" and previous actual values, $Y^*_t - Y_{t-1}$, and a straightforward response to ΔX_t. This is ADL(1,1) model. The ADL(1,0) model is a special case with the testable restriction $\alpha_4 = -\delta = 0$.