Lecture 17

Modeling with Time Series Data. Dynamic Processes.

Econometric models often include variables which are not directly observable. For example, an important role in economic analysis belongs to expectations (inflationary expectations, expectations of an exchange rate, etc.). However, there are no statistical data on expectations, because they are unobservable. There are many other unobservable economic variables: “permanent income” in M. Friedman’s model of consumption, “the target dividend rate” in G. Lintner’s model of dividend policy etc.

A common way of making a model with unobservable variables suitable for estimation and use in economic analysis is to introduce additional premises concerning the behaviour of unobservable variables and describe their interaction with observable ones. Two types of such premises will be considered in the lecture: a premise about adaptive expectations and a premise about partial adjustment. The main features of both models are given in the table below:

	Partial adjustment model	**Adaptive expectation model**
Unobserved variable:	**Dependent** variable	**Explanatory** variable
Adjustment process of:	Observed variable (actual level)	Unobserved variable
Model description:	$Y_t^* = \gamma_1 + \gamma_2 X_t + u_t$ $Y_t - Y_{t-1} = \lambda(Y_t^* - Y_{t-1})$ <=> where $Y_t = \lambda Y_t^* + (1-\lambda) Y_{t-1}$ Y_t^* – unobserved dependent variable	$Y_t = \gamma_1 + \gamma_2 X_{t+1}^e + u_t$ $X_{t+1}^e - X_t^e = \lambda(X_t - X_t^e)$ <=> where $X_{t+1}^e = \lambda X_t + (1-\lambda) X_t^e$ X_{t+1}^e – unobserved explanatory variable

They are identical from the mathematical point of view, since both include the calculation of a weighted average; the difference consists in the observability of the dependent and explanatory variables. In each models there are problems with parameter estimation, which will be discussed in the lecture.

I. Partial adjustment model:

Suppose, the unobservable (“target” or desired) value of the dependent variable is determined by the equation:

$$\boxed{Y_t^* = \beta_1 + \beta_2 X_t + u_t} \qquad (1)$$

However, this “target” value is reached gradually through a process of partial adjustment. There can be several ways to specify it. The common feature is that for non-observable dependent variable Y_t^*, inertia in the system makes the actual value of Y_t be a compromise (weighted average) between its value in the previous time period, Y_{t-1}, and the value of the unobserved variable justified by the value of the explanatory variable. The adjustment process can vary depending on the time period to which Y^* relates: either current value, Y_t^*, or one period lagged, Y_{t-1}^*. Let’s consider 2 cases:

1) *Adjustment to the current value of unobserved variable:* the actual increase in the dependent variable from time $t-1$ to time t, $Y_t - Y_{t-1}$, is proportional to the difference between the desired value of the dependent variable at the current period and its previous actual value, $Y_t^* - Y_{t-1}$: $Y_t - Y_{t-1} = \lambda(Y_t^* - Y_{t-1})$.

   This can be transformed to $Y_t = \lambda Y_t^* + (1-\lambda) Y_{t-1}$ where λ is called the speed of adjustment. λ lies in the range from 0 to 1. $\lambda = 0$ corresponds to no change at all, while $\lambda = 1$ means the full adjustment in the current period. Substituting for Y_t^* from the specification (1) results in the specification in terms of observable variables of the ADL(1,0) form:

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

where $\boxed{\alpha_1 = \beta_1\lambda, \quad \alpha_2 = \beta_2\lambda, \quad \alpha_3 = (1-\lambda)}$.

In the resulting model the explanatory variable and the disturbance term are not simultaneously correlated, therefore, the estimators will be consistent. However, OLS gives biased estimates as part (2) assumption of C.7 is now violated:

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

$Y_{t+1} = \alpha_1 + \alpha_2 X_{t+1} + \alpha_3 Y_t + \lambda u_{t+1}$ It is evident that u_t for Y_t affects the regressor Y_t for the next period dependent variable Y_{t+1}.

Short-run effect of X on Y: it is measured by α_2 coefficient. It equals $\beta_2\lambda$:

Long-run effect of X on Y: it is measured by β_2.

In fact, the relationship between the equilibrium values is written as:

$$\bar{Y} = \beta_1\lambda + \beta_2\lambda\bar{X} + (1-\lambda)\bar{Y}$$

Rearranging, getting how $\bar{Y}$ depends on $\bar{X}$:

$\lambda\bar{Y} = \beta_1\lambda + \beta_2\lambda\bar{X}$. Therefore, $\bar{Y} = \beta_1 + \beta_2\bar{X}$.

2) *Adjustment to the previous value of unobserved variable:* the actual increase in the dependent variable from time $t-1$ to time t, $Y_t - Y_{t-1}$, is proportional to the difference between the previous desired value of the dependent variable and its previous actual value, $Y_{t-1}^* - Y_{t-1}$: $Y_t - Y_{t-1} = \lambda(Y_{t-1}^* - Y_{t-1})$.

   This can be transformed to $Y_t = \lambda Y_{t-1}^* + (1-\lambda)Y_{t-1}$ Substituting for Y_{t-1}^* from the lagged by one period specification (1) results in the specification in terms of observable variables of the ADL(1,1) form:

$$\boxed{\begin{aligned} Y_t &= \lambda(\beta_1 + \beta_2 X_{t-1} + u_{t-1}) + (1-\lambda)Y_{t-1} \\ &= \beta_1\lambda + \beta_2\lambda X_{t-1} + (1-\lambda)Y_{t-1} + \lambda u_{t-1} \\ &= \alpha_1 + \alpha_2 X_{t-1} + \alpha_3 Y_{t-1} + \lambda u_{t-1} \end{aligned}}$$

   where $\boxed{\alpha_1 = \beta_1\lambda, \quad \alpha_2 = \beta_2\lambda, \quad \alpha_3 = (1-\lambda)}$.

   For the same reasons as in 1), results are biased but consistent.

   Short-run effect of *X* on *Y*: It equals zero. The dependent variable is not explained by the current value of X.

   Long-run effect of *X* on *Y*: It is measured by β_2 (the same derivation as before).

Example:

Lintner's model of dividend adjustment (1956):

In this model it is assumed that a firm has a desired level of dividends based on its expected earnings. When earnings (profit) vary, the firm will adjust its dividends slowly spreading these variations in earnings over a number of time periods. Lintner's model reflects empirically observed phenomenon that shareholders prefer smoothened dividend income. Suppose, D_t actual dividend, Π_t is profit, and D_t^* is "target" dividend. According to Lintner, the "target" dividend can be best explained by earnings to which the actual dividend is adjusting gradually through the partial adjustment process:

$$\boldsymbol{D_t^*} = \boldsymbol{\alpha} + \boldsymbol{\gamma\Pi}_t + u_t$$

$$D_t - D_{t-1} = \lambda(D_t^* - D_{t-1}) + v_t$$

Plugging the expression for the “target” dividend into the equation of partial adjustment, we get: $D_t - D_{t-1} = \lambda\alpha + \gamma\lambda\boldsymbol{\Pi}_t - \lambda D_{t-1} + \lambda u_t + v_t$

Accordingly, $\boxed{D_t = \lambda\alpha + \gamma\lambda\boldsymbol{\Pi}_t + (1-\lambda)D_{t-1} + \lambda u_t + v_t}$

In this model the explanatory variable $\boldsymbol{D_{t-1}}$and the disturbance term are related, but not simultaneously correlated, therefore, the estimators obtained will be biased, but consistent. The model can be directly estimated, since it does not include unobservable variables. G. Lintner has estimated the model on the data for the US corporate sector for 1918-1941 and has obtained the following results: $\boldsymbol{\gamma}$ *=0.3*; $\boldsymbol{\lambda}$ *=0.5*.

II. Adaptive expectations

There are 2 sources of dynamics for this type of models: inertia that is the drag from the past (like in the partial adjustment models) and the effect of anticipations. Economic agents form expectations about the future values of variables and then adjust their plans accordingly. For example, inflationary expectations influence the current interest rate and the demand for money, and the expected exchange rate affects the supply and demand for a currency. To estimate such type of models, it is necessary to introduce an assumption concerning the behaviour of the unobservable variable. In its simplest form, the dependent variable Y_t is explained by the anticipated value of X_{t+1}^e in the next period:

$$\boxed{Y_t = \beta_1 + \beta_2 X_{t+1}^e + u_t} \qquad (2)$$

X_{t+1}^e is subjective, so it is assumed to be described as the following process:

$$\boldsymbol{X_{t+1}^e - X_t^e = \lambda(X_t - X_t^e)}$$

λ is interpreted as a speed of adjustment (adaptation). It shows the share of expectations that are based on the actual behavior of the explanatory variable X, while $1-\lambda$ of them are related to the previous period anticipated value.

The adaptive expectation process can be rewritten as: $X_{t+1}^e = \lambda X_t + (1-\lambda)X_t^e$ (*)

Substituting for X_{t+1}^e from (2), we can obtain: $Y_t = \beta_1 + \beta_2\lambda X_t + \beta_2(1-\lambda)X_t^e + u_t$.

This equation still involves the unobserved variable. The adaptive expectation relationship (*) also holds for $t-1$: $X_t^e = \lambda X_{t-1} + (1-\lambda)X_{t-1}^e$.

Substituting for X_t^e in the equation for Y_t:

$$\boldsymbol{Y_t = \beta_1 + \beta_2\lambda X_t + \beta_2\lambda(1-\lambda)X_{t-1} + \beta_2(1-\lambda)^2 X_{t-1}^e + u_t}$$

Lagging and substituting s times in this way, we get:

$$\begin{aligned} Y_t = \beta_1 + \beta_2\lambda X_t + \beta_2\lambda(1-\lambda)X_{t-1} + \beta_2\lambda(1-\lambda)^2 X_{t-2} + \ldots \\ + \beta_2\lambda(1-\lambda)^{s-1} X_{t-s+1} + \beta_2(1-\lambda)^s X_{t-s+1}^e + u_t \end{aligned} \qquad (**)$$

For $\boldsymbol{s} \to \infty$, $(\boldsymbol{1-\lambda})^{\boldsymbol{s}}$ tends to zero => we can drop the unobservable final term. The specification is non-linear in parameters; therefore it can be fitted using some nonlinear estimation technique.

As $\boldsymbol{0 \le \lambda \le 1}$, it follows that $\boldsymbol{0 \le 1-\lambda \le 1}$ => we get a lag structure with geometrically declining weights – Koyck distribution. Using Koyck transformation the model can be expressed in terms of finite number of observable variables. Lagging (**) and then multiplying it by $\boldsymbol{1-\lambda}$,

$$(1-\lambda)Y_{t-1} = \beta_1(1-\lambda) + \beta_2\lambda(1-\lambda)X_{t-1} + \beta_2\lambda(1-\lambda)^2 X_{t-2} + \beta_2\lambda(1-\lambda)^3 X_{t-3} + ... + \beta_2\lambda(1-\lambda)^{s-1} X_{t-s+1} + \beta_2(1-\lambda)^s X^e_{t-s+1} + (1-\lambda)u_{t-1}$$

Subtracting the derived expression from (**) under $s \to \infty$ (=> final terms can be dropped):

$$\boldsymbol{Y_t - (1-\lambda)Y_{t-1} = \beta_1 + \beta_2\lambda X_t - \beta_1(1-\lambda) + u_t - (1-\lambda)u_{t-1}}$$

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

where $\boxed{\boldsymbol{\alpha_1 = \beta_1\lambda, \quad \alpha_2 = \beta_2\lambda, \quad \alpha_3 = (1-\lambda)}}$

Short-run effect of X on Y: it is measured by α_2 coefficient. It equals $\beta_2\lambda$.
Long-run effect of X on Y: it is measured by $\boldsymbol{\beta_2}$.

<u>Another possible way</u> to derive this relationship is to use the original specification one period lagged: $\boldsymbol{Y_{t-1} = \beta_1 + \beta_2 X^e_t + u_{t-1}}$. From this let's obtain the expression of $\boldsymbol{\beta_2 X^e_t}$:

$$\beta_2 X^e_t = Y_{t-1} - \beta_1 - u_{t-1}$$

Substituting for $\boldsymbol{\beta_2 X^e_t}$ in $\boldsymbol{Y_t = \beta_1 + \beta_2\lambda X_t + \beta_2(1-\lambda)X^e_t + u_t}$ (for $\boldsymbol{s = 1}$ in (**)):

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

<u>*Estimation:*</u>

Note that, mathematically it is the same ADL(1,0) model as for the partial adjustment process apart from the compounded disturbance term. Hence, it would be difficult to differentiate between these two models in practice, despite the fact that the approaches are opposite in spirit. This is an example of observational equivalence of two theories. However, the compounded disturbance term plays important role here: OLS estimates become biased and inconsistent because both parts of C.7 are now violated. Part (1): $\boldsymbol{u_{t-1}}$ affects $\boldsymbol{Y_{t-1}}$ in the observation for $\boldsymbol{Y_t}$. Part (2) the same reasons as for the partial adjustment process.

Therefore, the model has to be estimated step by step as the geometrically distributed lag model until the coefficients cease to differ by some specified number in 2 consecutive steps, i.e.

0) Use (**): $$\begin{aligned} Y_t &= \beta_1 + \beta_2\lambda X_t + \beta_2\lambda(1-\lambda)X_{t-1} + \beta_2(1-\lambda)^2 X_{t-2} + ... \\ &+ \beta_2(1-\lambda)^{s-1} X_{t-s+1} + \beta_2(1-\lambda)^s X^e_{t-s+1} + u_t \end{aligned};$$
1) Estimate the model for $\boldsymbol{s = 1}$;
2) Estimate the model for $\boldsymbol{s = 2}$;

……………………………..

$\boldsymbol{s'}$) Estimate the model for $\boldsymbol{s = s'}$ until fitted coefficients coincide with corresponding coefficients from the step $\boldsymbol{s' - 1}$) with the specified accuracy. Hence, choose estimation for which $\boldsymbol{s = s' - 1}$.

<u>*Example:*</u>

Cagan's model of hyperinflation

The basic equation of the model is the equation of the demand for real money balances as a function of the next period expected rate of inflation: $\boldsymbol{\frac{M}{P} = f(\pi^e_{t+1})}$.

In the conditions of hyperinflation, the appropriate time unit is normally a month or even a week. The standard Cagan's model is specified as follows:

$$\log\left(\boldsymbol{\frac{M}{P}}\right) = \boldsymbol{\alpha - \rho \cdot \pi^e_{t+1} + u_t}$$

The parameter $\boldsymbol{\rho}$ in this model shows the percentage decrease in the demand for real money balances associated with a one percent increase in the expected inflation rate.
Let $\boldsymbol{\pi^e_{t+1}}$ follow a partial adjustment process:

$$\pi^e_{t+1} - \pi^e_t = \beta(\pi_t - \pi^e_t)$$

Then, $\log(\frac{M}{P})_t = \alpha - \rho\beta(\pi_t + (1-\beta)\pi_{t-1} + (1-\beta)^2\pi_{t-2} + \ldots) + u_t$

Using monthly data for 7 cases of a hyperinflation, he has obtained the following estimates of the model parameters: $\boldsymbol{\rho}$ = 4.68; $\boldsymbol{\beta}$ = 0.20. They indicate that in a hyperinflation the demand for real money balances is not very sensitive to the expected inflation rate, and the inflationary expectations are adjusted rather slowly.

Friedman's Permanent Income Hypothesis:

Lecture 10 considered Milton Friedman's Permanent Income Hypothesis (PIH) as the application of the measurement error analysis. According to this theory, permanent consumption is proportional to permanent income: $\boldsymbol{C^P_t = \beta_2 Y^P_t}$. Permanent income is a subjective concept of likely medium-run future income. Actual income and consumption is decomposed into permanent and transitory components:

$$\boldsymbol{Y_t = Y^P_t + Y^T_t}$$
$$\boldsymbol{C_t = C^P_t + C^T_t}$$

Since the permanent income is unobservable, suppose that it follows the adaptive expectation process: $\boldsymbol{Y^P_t - Y^P_{t-1} = \lambda(Y_t - Y^P_{t-1})}$. This can be rewritten as follows: permanent income at $\boldsymbol{t}$ is a weighted average of actual income at $\boldsymbol{t}$ and permanent income at $\boldsymbol{t-1}$, i.e. $\boldsymbol{Y^P_t = \lambda Y_t + (1-\lambda)Y^P_{t-1}}$. This can be substituted into the consumption function for $\boldsymbol{Y^P_t}$:

$$\boldsymbol{C_t - C^T_t = \beta_2\big(\lambda Y_t + (1-\lambda)Y^P_{t-1}\big)}$$

By lagging the adaptive process $\boldsymbol{Y^P_{t-1} = \lambda Y_{t-1} + (1-\lambda)Y^P_{t-2}}$ and substituting it for $\boldsymbol{Y^P_{t-1}}$, we get:

$$\boldsymbol{C_t = \beta_2\lambda Y_t + \beta_2\lambda(1-\lambda)Y_{t-1} + \beta_2(1-\lambda)^2 Y^P_{t-2} + C^T_t}$$

Repeating this procedure $\boldsymbol{s}$ times, we obtain:

$$\boldsymbol{C_t = \beta_2\lambda Y_t + \beta_2\lambda(1-\lambda)Y_{t-1} + \beta_2\lambda(1-\lambda)^2 Y_{t-2} + \ldots}$$
$$\boldsymbol{+\beta_2\lambda(1-\lambda)^{s-1}Y_{t-s+1} + \beta_2(1-\lambda)^s Y^P_{t-s} + C^T_t}$$

It is non-linear in parameters specification. Friedman fitted it using non-linear iterative estimation method. In order to evaluate short-run and long-run effects, we can either use Koyck transformation or the described in the previous section another way to get rid of unobservable variables. Lagging the basic relationship one period $\boldsymbol{\beta_2 Y^P_{t-1} = C^P_{t-1} = C_{t-1} - C^T_{t-1}}$. Substituting it into $\boldsymbol{C_t - C^T_t = \beta_2\big(\lambda Y_t + (1-\lambda)Y^P_{t-1}\big)}$, we obtain:

$$C_t = \lambda\beta_2 Y_t + (1-\lambda)(C_{t-1} - C^T_{t-1}) + C^T_t$$
$$= \lambda\beta_2 Y_t + (1-\lambda)C_{t-1} + C^T_t - (1-\lambda)C^T_{t-1}$$

Short run marginal propensity to consume equals $\boldsymbol{mpc_{SR} = \lambda\beta_2}$;
Long run marginal propensity to consume equals $\boldsymbol{mpc_{LR} = \beta_2}$.

Thus, since $\boldsymbol{\lambda}$ is less than 1, the model is able to explain the fact that after the Second World War the long-run average propensity to consume seemed to be roughly constant despite the marginal propensity to consume being much lower.