ECON 4160, Spring term 2013. Lecture 3

GLS. Concepts of dynamic modeling

Ragnar Nymoen

University of Oslo

4 September 2013

Some references to Lecture 3

- DM Ch 7.1–7.5, (Generalized least squares)
- DM Ch 7.6 and 13.1 (Lag-operators and Stationarity)
- DM CH 13.4 (ADL model)

GLS I

- We know that both heteroskedasticty and autocorrelation require a different specification than $Var(\varepsilon) = E(\varepsilon\varepsilon^{'}) = \sigma^2\mathbf{I}$ in the linear regression model

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \tag{1}$$

More generally we have:

$$Var(\varepsilon) = \sigma^2\Omega,\ \sigma^2 > 0$$

where Ω is $n \times n$ is symmetric and **Positive Definite:**

$$\underbrace{\mathbf{z}'\Omega\mathbf{z}}_{\text{quadratic form in } n \text{ variables}} > 0 \text{ for all } \mathbf{z} \neq \mathbf{0}$$

with $\mathbf{z}' = (Z_1, Z_2, \ldots, Z_n)$.

GLS II

- Result from linear algebra: For a PD matrix $\mathbf{\Omega}$ there exists a $n \times n$ matrix $\mathbf{\Psi}$ which is invertible (non-singular), with properties

$$\mathbf{\Psi}\mathbf{\Omega}\mathbf{\Psi}' = \mathbf{I} \tag{2}$$

$$\mathbf{\Psi}^{'}\mathbf{\Psi} = \mathbf{\Omega}^{-1} \Leftrightarrow \mathbf{\Psi}\mathbf{\Psi}' = \mathbf{\Omega}^{-1} \text{ (symmetry of } \mathbf{\Omega}) \tag{3}$$

- Multiplication from the left in (1) by $\mathbf{\Psi}^{'}$ gives:

$$\underbrace{\mathbf{\Psi}^{'}\mathbf{Y}}_{\mathbf{y}_*} = \underbrace{\mathbf{\Psi}^{'}\mathbf{X}}_{\mathbf{X}_*}\beta + \underbrace{\mathbf{\Psi}^{'}\varepsilon}_{\varepsilon_*} \tag{4}$$

GLS III

- Because

$$Var(\varepsilon_*) = E(\varepsilon_* \varepsilon_*^{'}) = E(\mathbf{\Psi}^{'} \varepsilon \varepsilon^{'} \mathbf{\Psi}) = \mathbf{\Psi}^{'} \sigma^2 \mathbf{\Omega} \mathbf{\Psi} = \sigma^2 \mathbf{I} \quad (5)$$

  the OLS estimator for β from (4) is BLUE under the assumption of strict exogeneity of the X-variables.
- This estimator is the **Generalized Least Squares** estimator (GLS) and is, by reference to minimization of residuals and to method-of-moments, given by:

$$\hat{\beta}_{GLS} = (\mathbf{X}_*^{'} \mathbf{X}_*)^{-1} \mathbf{X}_*^{'} \mathbf{y}_* = (\mathbf{X}' \mathbf{\Psi} \mathbf{\Psi}^{'} \mathbf{X})^{-1} \mathbf{X}^{'} \mathbf{\Psi} \mathbf{\Psi}^{'} \mathbf{y}$$
$$= (\mathbf{X}' \mathbf{\Omega}^{-1} \mathbf{X})^{-1} \mathbf{X}^{'} \mathbf{\Omega}^{-1} \mathbf{y} \quad (6)$$

  with covariance matrix.

$$Var(\hat{\beta}_{GLS}) = (\mathbf{X}_*^{'} \sigma^{-2} \mathbf{I} \mathbf{X}_*)^{-1} = \sigma^2 (\mathbf{X}' \mathbf{\Omega}^{-1} \mathbf{X})^{-1}$$

Weighted Least Squares example I

Assume that the only departure from the classical assumptions is heteroskedasticity, and that it takes the form:

$$Var(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{\Omega} = \sigma^2 \begin{pmatrix} X_{21} & 0 & \cdots & 0 \\ 0 & X_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_{2n} \end{pmatrix}.$$

Then (show!):

$$\boldsymbol{\Omega}^{-1} = \begin{pmatrix} \frac{1}{X_{21}} & 0 & \cdots & 0 \\ 0 & \frac{1}{X_{21}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{X_{2n}} \end{pmatrix}$$

Weighted Least Squares example II

and we can compute the GLS estimator from (6).
Moreover, we see that $\mathbf{\Psi\Omega\Psi}' = \mathbf{I}$ if $\mathbf{\Psi} = \mathbf{\Psi}^{'}$ is specified as :

$$\mathbf{\Psi}' = \begin{pmatrix} \sqrt{\frac{1}{X_{21}}} & 0 & \cdots & 0 \\ 0 & \sqrt{\frac{1}{X_{21}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\frac{1}{X_{2n}}} \end{pmatrix}$$

which gives the weights that we apply to obtain $\mathbf{y}_*$ and $\mathbf{X}_*$ in (4).

Feasible GLS I

- In most practical situations $\boldsymbol{\Omega}$, is unknown as is replaced by a consistent estimator $\hat{\boldsymbol{\Omega}}$. This is based on
  - the OLS residuals $\hat{\varepsilon}_i$
  - and an **assumed form** of the heteroskedasticity or autocorrelation (see DM section 7.4 for examples).
- In practice: an auxiliary regression between $\hat{\varepsilon}_i$ and a set of observable variables (often some of the X variables).
- As long as this procedure gives a consistent estimator of $\boldsymbol{\Omega}$, the **feasible** GLS estimator

$$\tilde{\beta}_{GLS} = (\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{y}$$

  is both **consistent** and **asymptotically efficient**.

Conditional and marginal models I

- A model of the simultaneous pdf $f(Y, X_1, X_2,\ldots,X_k)$ is an econometric simultaneous equations model.
- That simultaneous equation model is one of the main topic later in the course.
- First we will however wee how we can represent a system of variables with what we can loosely call, **a system of regression models**.
- The conditional of Y model based on the conditional pdf $f(Y \mid X_1, \ldots, X_k)$ is a regression model.

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \tag{7}$$

where we assume linearity and that ε have the classical properties for simplicity

Conditional and marginal models II

- Clearly (7) is only a partial model of the system.
- Can "close the model" if we supplement (7) with equations that represent the marginal pdf $f(X_1, \ldots, X_k)$.
- If the variables in $\mathbf{X}$ are IID, the marginal model will be the k (super) simple regressions:

$$\mathbf{X}_j = \boldsymbol{\phi}_j + \boldsymbol{v}_j \; j = 1, 2, \ldots, k \tag{8}$$

where $\boldsymbol{\phi}_j$ ($j = 1, 2, \ldots, k$) are parameters and

$$E(\boldsymbol{v}_j) = \mathbf{0}, \; , \forall j$$

$$Var(\boldsymbol{v}_j) = E(\boldsymbol{v}_j \boldsymbol{v}_j^{'}) = \sigma_v^2 \mathbf{I}, \; \forall j$$

$$E(\boldsymbol{v}_j \boldsymbol{v}_l) = 0 \; , \; \forall j \neq l$$

Conditional and marginal models III

- (7) and (8) give a model representation of the joint pdf $f(Y,$ $X_1, X_2, \ldots, X_k)$ for the case of IID regressors.
- The assumption of IID regressors is often unrealistic in particular for time series data.
- Hoever,there are purposes that require us to model the marginal model, for example if the purpose is to forecast Y one or more periods ahead.
- In the rest of the course we therefore turn to **time series data**, and to **dynamic** equations and systems.
- As a stepping stone, we need some basic concepts and theorems from time series econometrics, before we return to model specification and estimation (CC 2 and Lecture4)

Time series I

- We define a time series Y_t as the realization of a *stochastic process* $\{Y_t; t \epsilon T\}$. In any period t the variable Y_t can take a number of values consistent with the the sample space. (Norwegian: "utfallsrom").
- A stochastic process has therefore a random distribution for each Y_t. It is consistent with this definition that T can be $\{0, \pm 1, \pm 2, \ldots\}$, $\{1, 2, 3, \ldots\}$, $[0, \infty\}$ or $(-\infty, \infty)$.
- However, when there is no room for misunderstanding, we follow convention and use the term *time series* both for a data series, and for the process of which it is a realization.
- We will mainly study stochastic processes given by **linear stochastic difference equations**.

A time series of order p, (AR(p) I

- We write a time series of order p as the stochastic difference equation

$$Y_t = \gamma + \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_p Y_{t-p} + \varepsilon_t \qquad (9)$$

where $\gamma, \rho_j \ (j = 1, 2, \ldots, p)$ are parameters, and where

$$\varepsilon_t \sim IID\left(0, \sigma_\varepsilon^2\right) \quad \forall \ t. \qquad (10)$$

as in equation (13.01) in DM who refer to this ε_t as **white-noise**.

- Together they are known as the AR(p) model.
- (9) may of interest "on its own", as a general model of single time series.

A time series of order p, (AR(p) II

- One example is when Y_t is not a an observable variable, but a residual from OLS estimation.
- In that interpretation (9) becomes a model of autocorrelated regression residuals, as covered in Ch 7.6 in DM.
- Estimate by NLS or feasible GLS, possibly iterated.
- When Y_t is an observable, the main motivation for using (9) is for forecasting.
- The reason for studying (9) in econometics is however, more fundamental: It gives the framework for defining the all important concepts of **dynamic stability** and **stationarity** both for individual time series and for systems of variables (for example dynamic stochastic general equilibirum models,DSGE).

Prelude: AR(p) as the final equation of a system I

- We ofen study systems of stochastic difference equations
- The simplest case is two time series that are connected in a the first order system .

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix}, \qquad (11)$$

where $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is the matrix of autoregressive coefficients and we assume that

$$\begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix} \sim IID \left(\mathbf{0}, \begin{matrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{matrix} \right) \forall\ t \qquad (12)$$

Prelude: AR(p) as the final equation of a system II

- In fact this is an example of a first order Vector Autoregressive model, **VAR**.
- If (12) is replaced by the normal (Gaussian) distribution, the system is called a **Gaussian VAR.**
- As an exercise, you can show that (11) can be reduced to the so called **final equation** for Y_{t+1}

$$Y_{t+1} = (\underbrace{a_{11} + a_{22}}_{\equiv \rho_1})Y_t + (\underbrace{a_{12}a_{21} - a_{22}a_{11}}_{\equiv \rho_2})Y_t + \underbrace{\varepsilon_{yt+1} - a_{22}\varepsilon_{yt} + a_{12}\varepsilon_{}}_{\equiv \varepsilon_t} \quad (13)$$

- The answer will be posted on web page after the lecture.

Prelude: AR(p) as the final equation of a system III

- But the same equation must hold for Y_t so we obtain (9) for the case of $p = 2$ and $\gamma = 0$ as

$$Y_t = \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \varepsilon_t \tag{14}$$

$$\rho_1 = (a_{11} + a_{22}) \tag{15}$$

$$\rho_2 = a_{12}a_{21} - a_{22}a_{11} \tag{16}$$

$$\varepsilon_t = \varepsilon_{y,t} - a_{22}\varepsilon_{y,t-1} + a_{12}\varepsilon_{x,t-1} \tag{17}$$

- The omission of the intercept (which implies $\gamma = 0$) is only to save notation.

Prelude: AR(p) as the final equation of a system IV

- Note that when ε_t is defined as in (17) we have $E(\varepsilon_t) = 0$ and

$$\begin{aligned} Var(\varepsilon_t) &= Var(\varepsilon_{y,t} - a_{22}\varepsilon_{y,t-1} + a_{12}\varepsilon_{x,t-1}) \\ &= \sigma_y^2 + a_{22}^2\sigma_{yy} + a_{12}^2\sigma_x^2 + 2a_{22}a_{12}\sigma_{yx} \end{aligned}$$

  independent of t (homoskedasticity), but

$$\begin{aligned} Cov(\varepsilon_t, \varepsilon_{t-1}) &= -a_{22}\sigma_y^2 + a_{12}\sigma_{yx} \\ Cov(\varepsilon_t, \varepsilon_{t-j}) &= 0 \; j = 2, 3, \ldots \end{aligned}$$

- In this interpretation the disturbance ε_t in the AR(2) model is not white-noise, but a **Moving Average** (MA) process

Dynamic stability and stationarity I

- Consider again the AR(p) process:

$$Y_t = \gamma + \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + ... + \rho_p Y_{t-p} + \varepsilon_t \qquad (18)$$

- Consider next the **homogenous version** of the difference equation:

$$Y_t^h - \rho_1 Y_{t-1}^h - \rho_2 Y_{t-2}^h - ... - \rho_p Y_{t-p}^h = 0 \qquad (19)$$

Dynamic stability and stationarity II

- From mathematics we know that (19) has a **global asymptotic stable solution** ($Y_t^h \to 0$ when $t \to \infty$) if and only if all the p roots (eigenvalues) of the associated characteristic polynomial

$$\lambda^p - \rho_1 \lambda^{p-1} - \rho_2 \lambda^{p-2} - ... - \rho_p = 0 \qquad (20)$$

  are less than one in absolute value.
- From a result that is far from trivial, and which we leave for ECON 5101, we have that the same condition is necessary and sufficient for the **stationarity** of the stochastic process Y_t when it is given by (18) and ε_t is white-noise, or any other stationary time series process (e.g., MA(q), $q = 1, 2, ...$).
- But now we have given the condition for stationarity without a definition for stationarity...!

Stationarity defined I

For the time series $\{Y_t;\ t = 0, \pm 1, \pm 2, \pm 3, ...\}$ we define the *autocovariances* $\gamma_{j,t}$ as

$$\tau_{j,t} = E[(Y_t - \mu_t)(Y_{t-j} - \mu_t)],\ \ j = 0, 1, 2, \ldots, \tag{21}$$

where$E(Y_t) = \mu_t$.
If neither μ nor $\gamma_{j,}$ depend on time t:

$$E(Y_t) = \mu, \forall\ t$$

and

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = \tau_j, \forall\ t,\ j.$$

the Y_t process $\{Y_t;\ t = 0, \pm 1, \pm 2, \pm 3, ...\}$ is **covariance stationary** (aka weakly stationary).

Stationarity defined II

For a stationary Y_t the variance is time independent

$$Var(Y_t) = \sigma_y^2 \equiv \tau_0 \text{ for } j = 0$$

and the autocovariances are symmetric backwards and forwards: $\tau_j = \tau_{-j}$

The autocorrelation function I

- For a stationary time series variable, the theoretical autocovariances only depend on the distance j between periods. We can regard the autocovariance as a function of j.
- The same is the case for the (theoretical) autocorrelation function (ACF). In general, it is a function of j and t:

$$\zeta_{j,t} = \{Y_t, Y_{t-j}\} = \frac{Cov(Y_t, Y_{t-j})}{Var(Y_t)} = \frac{\tau_{j,t}}{\tau_{0,t}}, \qquad (22)$$

However

$$\zeta_j = \frac{\tau_j}{\tau_0} = \zeta_{-j} \text{ for } j = 1, 2, ... \qquad (23)$$

in the stationary case.

Why is stationarity so important? I

- For an observable time series $\{Y_t;\ t = 1, 2, 3, ... T\}$, we use the empirical autocovariances,

$$\hat{\tau}_j = 1/T \sum_{t=j+1}^{T} (Y_t - \bar{Y})(Y_{t-j} - \bar{Y}),\ j = 0, 1, 2, \ldots, T-1 \tag{24}$$

  where $\bar{Y} = 1/T \sum_{t=1}^{T} Y_t$.
- If the process $\{Y_t;\ t = 0, \pm 1, \pm 2, \pm 3, ...\}$ is stationary, $\hat{\tau}_j$ $(j = 0, 1, 2, \ldots)$ are consistent estimators of the theoretical autocovariances
- This in turn gives the main premise for consistent estimation of the coefficients of dynamic regression models, of which AR(p) is an example

Why is stationarity so important? II

- In short: stationary is the main premise for why we can extend the OLS based estimation and inference theory to time series data!
- Note that, although stationarity depends on the characteristics roots, it can be "mapped back" to the ρ_1 and ρ_2 parametes in the AR(2) case. See Figure 13.1 in DM

AR(2) example I

$\gamma = 0,\ \ \rho_1 = 1,6,\ \rho_2 = -0,9$:

$$Y_t = 1,6Y_{t-1} - 0,9Y_{t-2} + \varepsilon_t, \tag{25}$$

The roots are a complex pair. The “absolute value” of the roots is 0.94868.

- Show homogenous solution,
- and solution when $\varepsilon_t \sim IID(0,1)$ in class

Lag operators I

- When we work with stochastic difference equations, it is often useful to express relationships with the use of the lag-operator L.
- The lag operator L changes the dating of a variable Y_t one or more period back in time. It works in the following way:

$$\begin{aligned} LY_t &= Y_{t-1}, \\ LLY_t = L^2 Y_t &= LY_{t-1} = Y_{t-2}, \\ L^p Y_t &= Y_{t-p}. \end{aligned}$$

- From the last property it follows that if $p = 0$, then

$$\begin{aligned} L^0 &= 1, \\ L^0 Y_t &= Y_t. \end{aligned}$$

Lag operators II

- We also have
$$L^p L^s = L^p L^k = L^{(p+s)},$$
and
$$\left(aL^p + bL^s\right) Y_t = aL^p Y_t + bL^s Y_t = aY_{t-p} + bY_{t-s}.$$
- If we want to shift a variable forward in time, we use the forward operator L^{-1}:
$$L^{-1} Y_t = Y_{t+1}$$
and generally
$$L^{-s} = Y_{t+s}.$$

Lag operators III

- Because constants are independent of time, we have for the constant b

$$Lb = b.$$

and by induction

$$L^p b = L^{(p-1)} L b = L^{(p-1)} b = b.$$

Lag-polynomial representation of AR(p) I

- We can now write (9) more compactly as

$$\phi(L)Y_t = \gamma + \varepsilon_t \tag{26}$$

where is the lag polynomial of order p.

$$\phi(L)Y_t = 1 - \rho_1 L - \rho_2 L^2 - ...\rho_p L^p \tag{27}$$

and we keep the assumption of white-noise ε_t.

Lag-polynomial representation of AR(p) II

- A root of the characteristic equation associated with the lag-polynomial is:

$$1 - \rho_1 z - \rho_2 z - \ldots \rho_p z^p = 0 \tag{28}$$

Comparison with the characteristic equation (20) shows that

$$z = \frac{1}{\lambda}$$

meaning that the condition for stationarity can also be expressed in terms of the roots: $(z_1, z_2, \ldots, z_p)$:

- Y_t is stationary if all the $z-$roots are larger than one in absolute value ("outside the unit circle" DM page 273).

Companion form I

Consider again the VAR system (11)

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}_{A} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix},$$

- $(Y_t, X_t)'$ and $(\varepsilon_{yt}, \varepsilon_{xt})'$ are two bivariate time series.
- Assume that $(\varepsilon_{yt}, \varepsilon_{xt})'$ are made up of two stationary series. This is secured by (12) for example.

Companion form II

- By obtaining the characteristic polynomial to **A**:

$$p(\lambda) = |\mathbf{A} - \lambda \mathbf{I}|$$

you find that the **eigenvalues of A** are the roots of

$$|\mathbf{A} - \lambda \mathbf{I}| = 0 \tag{29}$$

which is the characteristic equation associated with the final equation (13) that we derived above.

- Hence the necessary and sufficient condition for stationary of the vector $(Y_t, X_t)'$ is that the two eigenvalues of both less than one in absolute value.
- **A** is a simple example of a so called **companion form** matrix.

Companion form III

- In ECON 5101 we will show that if we have a general VAR with n time series variables and p lags, that VAR can be written as a first order system

$$\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t-1} + \epsilon \tag{30}$$

where $\mathbf{z}_t$ and ϵ_t are $1 \times np$ and the companion form matrix $\mathbf{F}$ is $np \times np$.

- For such a general VAR system, the condition for stationarity and stability is that all the np eigenvalues from

$$|\mathbf{F} - \lambda \mathbf{I}| = 0 \tag{31}$$

are less than one in magnitude.

Companion form IV

- When we estimate a dynamic system in PcGive, the eigenvalues of the companion form are always available after estimation.

The VAR revisited I

Let us now take care to write the Gaussian disturbances of the VAR (now including two intercepts)

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \pi_{10} \\ \pi_{20} \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix} \tag{32}$$

as conditional on period $t-1$:

$$\begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \mid Y_{t-1}, X_{t-1}\right). \tag{33}$$

Now, (32) can be written as

$$Y_t = \mu_{y,t-1} + \varepsilon_{yt} \tag{34}$$

$$X_t = \mu_{x,t-1} + \varepsilon_{xt} \tag{35}$$

The VAR revisited II

where the **conditional expectations** $\mu_{y,t-1} \equiv E(Y_t \mid Y_{t-1}, X_{t-1})$ and $\mu_{x,t-1} \equiv E(X_t \mid Y_{t-1}, X_{t-1})$ are

$$\mu_{y,t-1} = \pi_{10} + \pi_{11} Y_{t-1} + \pi_{12} X_{t-1} \quad (36)$$

$$\mu_{x,t-1} = \pi_{20} + \pi_{21} Y_{t-1} + \pi_{22} X_{t-1}. \quad (37)$$

Interpretation: Conditional on the history of the system up to time $t-1$, Y_t and X_t are jointly normally distributed.

The conditional model for Y I

The conditional distribution for Y_t given the history **and** X_t is also normal,
In **Lecture note 3** (posted after the lecture for self-study) we show that the conditional distribution for Y_t is:

$$Y \sim N(\phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1}, \sigma^2 \mid X_t, Y_{t-1}, X_{t-1}) \quad (38)$$

which can be written in "model form" as

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \quad (39)$$

$$\varepsilon_t \sim N(0, \sigma^2 \mid X_t, Y_{t-1}, X_{t-1}) \quad (40)$$

The conditional model for Y II

$$\phi_0 = \pi_{10} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{20} \tag{41}$$

$$\phi_1 = \pi_{11} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{21} \tag{42}$$

$$\beta_0 = \frac{\sigma_{xy}}{\sigma_x^2} \tag{43}$$

$$\beta_1 = \pi_{12} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{22} \tag{44}$$

and

$$\sigma^2 = \sigma_y^2(1 - \rho_{xy}^2). \tag{45}$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}. \tag{46}$$

The conditional model for Y III

- Some small differences in notation apart, this is the same ADL model as in Ch 13.5 eq (13.58) for $p = q = 1$.
- The same ADL type model can be derived from a VAR with IID disturbances, rather than strictly normal.
- ADL(p,q) model can be derived from a VAR or order p. Consequently we must then have $p = q$ in the ADL.
- We will study such ADL models, and their estimation over the next weeks.

The conditional model for Y IV

- Finally, note that the ADL model

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \tag{47}$$

together with the second row in the VAR:

$$X_t = \pi_{20} + \pi_{21} Y_{t-1} + \pi_{22} X_{t-1} + \varepsilon_{xt} \tag{48}$$

give a **regression representation** of the VAR, in terms of a **conditional model** (47) and a **marginal model** (47).

- Meaning that we have extende the regression model representation of the static simultaneous system to the dynamic model case.