# AUTOCORRELATION

YUSEP SUPARMAN

# THE GAUSS-MARKOV MODEL

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

$$\mathrm{E}\left(\boldsymbol{\varepsilon} \middle| \mathbf{X}\right) = \mathrm{E}\left(\boldsymbol{\varepsilon}\right) = \mathbf{0}$$

$$\mathrm{cov}\left(\boldsymbol{\varepsilon} \middle| \mathbf{X}\right) = \mathrm{cov}\left(\boldsymbol{\varepsilon}\right) = \sigma^2 \mathbf{I}$$

AUTOCORRELATION CONDITION

$$\mathrm{cov}\left(\boldsymbol{\varepsilon} \middle| \mathbf{X}\right) = \begin{bmatrix} \sigma^2 & \sigma_{21} & \cdots & \sigma_{n1} \\ \sigma_{21} & \sigma^2 & \cdots & \sigma_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \mathbf{\Psi};\ \mathbf{\Psi} \neq \mathbf{I}$$

# THE AUTOCORRELATION EFFECT

$$\begin{aligned}
\operatorname{var}\left(\hat{\boldsymbol{\beta}} \middle| \mathbf{X}\right) &= \operatorname{var}\left\{\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \middle| \mathbf{X}\right\} \\
&= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\operatorname{var}\left(\boldsymbol{\varepsilon} \middle| \mathbf{X}\right)\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \\
&= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Psi}\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \\
&\neq \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}; \boldsymbol{\Psi} \neq \mathbf{I}
\end{aligned}$$

Wrong standard error – invalid inference

# FIRST ORDER TEMPORAL AUTOREGRESSION

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + \upsilon_t$$

$$\mathrm{E}(\upsilon_t) = 0 \qquad \mathrm{var}(\upsilon_t) = \sigma_\upsilon \qquad \mathrm{cov}(\upsilon_t, \upsilon_s) = 0$$

$$t, s = 1, 2, \ldots, T$$

Stationary condition $|\rho| < 1$

$$\rightarrow \mathrm{E}(\varepsilon_t) = 0, \mathrm{var}(\varepsilon_t) = \sigma_\varepsilon^2, \mathrm{cov}(\varepsilon_t, \varepsilon_{t-s}) = \sigma_s$$

# FIRST ORDER TEMPORAL AUTOREGRESSION

$$\mathrm{E}(\varepsilon_t) = \rho\,\mathrm{E}(\varepsilon_{t-1}) + \mathrm{E}(\upsilon_t) = 0 \leftarrow \mathrm{E}(\varepsilon_1) = \mathrm{E}(\upsilon_t) = 0$$

$$\mathrm{var}(\varepsilon_t) = \mathrm{var}(\rho\varepsilon_t + \upsilon_t) = \rho^2\,\mathrm{var}(\varepsilon_{t-1}) + \mathrm{var}(\upsilon_t) \leftarrow \mathrm{cov}(\varepsilon_t, \upsilon_s) = 0$$

$$= \frac{\sigma_\upsilon^2}{1-\rho^2} \leftarrow \mathrm{var}(\varepsilon_t) = \mathrm{var}(\varepsilon_{t-1}), \mathrm{var}(\upsilon_t) = \sigma_\upsilon^2$$

$$\mathrm{cov}(\varepsilon_t, \varepsilon_{t-1}) = E(\varepsilon_t\varepsilon_{t-1}) = \rho E(\varepsilon_{t-1}^2) + E(\varepsilon_{t-1}\upsilon_t) = \rho\frac{\sigma_\upsilon^2}{1-\rho^2}$$

$$E(\varepsilon_t\varepsilon_{t-2}) = \rho E(\varepsilon_{t-1}\varepsilon_{t-2}) + E(\varepsilon_{t-2}\upsilon_t) = \rho^2\frac{\sigma_\upsilon^2}{1-\rho^2}$$

$$E(\varepsilon_t\varepsilon_{t-s}) == \rho^s\frac{\sigma_\upsilon^2}{1-\rho^2}$$

# GLS ESTIMATOR

Upsilon is an error term meeting Gauss-Markov conditions. Accordingly we can use it as the error term in the following transformed model:

$$\varepsilon_t = \rho\varepsilon_{t-1} + \upsilon_t$$

$$y_t - \mathbf{x}_t'\beta = \rho\left(y_{t-1} - \mathbf{x}_{t-1}'\beta\right) + \upsilon_t$$

$$y_t - \rho y_{t-1} = \left(\mathbf{x}_t - \rho\mathbf{x}_{t-1}\right)'\beta + \upsilon_t$$

$$t = 2, 3, \ldots, T$$

OLS of the transformed model is only an approximation to GLS since, the first observation is excluded from the analysis (Cochrane-Orcutt estimator).

For a large sample size, the exclusion will not have a large impact on the results.

# GLS ESTIMATOR

Rescuing the first observation by assuming that **epsilon 1 is uncorrelated with all upsilon**

Since variance of epsilon 1 is larger than variance of upsilon, the first observation should be transformed by multiplying it by square root of one minus rho square, giving the following model

$$\sqrt{1-\rho^2}\, y_1 = \sqrt{1-\rho^2}\,\mathbf{x}_1'\boldsymbol{\beta} + \sqrt{1-\rho^2}\,\varepsilon_1$$

Applying OLS on the two models produces a BLUE GLS estimator (Prais-Winsten estimator).

# EGLS ESTIMATOR
# ITERATIVE COCHRANE-ORCUTT PROCEDURE

Using rho-hat instead of rho in the transformation.

Rho-hat from OLS is biased but consistent.

Accordingly an iterative procedure is required:

1. Apply OLS on untransformed data, calculate the residuals
2. Estimate rho-hat

$$\hat{\rho} = \left( \sum_{t=2}^{T} e_{t-2}^{2} \right)^{-1} \left( \sum_{t=2}^{T} e_t e_{t-1} \right)$$

3. Apply EGLS, obtain beta-hat.
4. Calculate residuals and rho-hat.
5. Repeat step 3-4 until beta-hat and rho-hat are convergent.

The estimator is not BLUE but consistent and efficient.

# HETEROSKEDASTICITY AND AUTOCORRELATION CONSISTENT STANDARD ERROR

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t$$

$$\mathrm{E}\left(\mathbf{x}_t \varepsilon_t\right) = 0 \qquad E\left(\varepsilon_t \varepsilon_{t-s}\right) = 0 \text{ for s=H,H+1,}\ldots$$

$$\hat{\mathrm{cov}}\left(\hat{\beta}_{OLS}\right) = \left(\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right)^{-1} TS^* \left(\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right)^{-1}$$

$$S^* = \frac{1}{T}\sum_{t=1}^{T} e_t^2 \mathbf{x}_t \mathbf{x}_t' + \frac{1}{T}\sum_{j=1}^{H-1} w_j \sum_{s=j+1}^{T} e_s e_{s-j}\left(\mathbf{x}_s \mathbf{x}_{s-j}' + \mathbf{x}_{s-j}\mathbf{x}_s'\right)$$

$$w_j = 1 - j/H$$