

# Econometrics

## Chapter 9: Basic Regression Analysis with Time Series

In Choi

Sogang University

# Time Series vs. Cross Sectional

- Time series data has a temporal ordering, unlike cross-section data.
- We no longer have an iid sample of individuals.
- Instead, we have one realization of a stochastic (i.e. random) process.

# Examples of Time Series Models

- A static model relates contemporaneous variables:

$$y_t = \alpha_0 + \delta_0 z_t + u_t.$$

## Example

### Static Phillips curve

$$infl_t = \beta_0 + \beta_1 unemp_t + u_t.$$

- A finite distributed lag (FDL) model allows one or more variables to affect  $y$  with a lag:

$$y_t = \alpha_0 + \delta_0 z_t + \dots + \delta_q z_{t-q} + u_t.$$

- Autoregressive model

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t.$$

- EViews tips for the AR model estimation:
  - Use command "series c series(-1)" for the AR(1) regression
  - Use the "Structure/resize" button under "Proc" before forecasting.

# Assumptions for Unbiasedness

- We need the same assumptions as in Chapter 3. That is,  $E(u_t \mid \text{all } x) = 0$  for all  $t$ .
- Note that the assumptions do not hold for the AR model.

**Illustration** Consider the AR(1) model

$$y_t = \alpha y_{t-1} + u_t, (t = 2, \dots, T).$$

The usual assumption in linear regressions is

$$E(u_t \mid \text{all } x's) = 0. \quad (1)$$

In the AR(1) model, this amounts to

$$E(u_t \mid y_2, \dots, y_T) = 0.$$

But

$$y_t = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \dots,$$

so that

$$E(u_t \mid y_2, \dots, y_T) = u_t.$$

The usual assumption (1) cannot hold for the AR(1) model.

# Unbiasedness of OLS

- Under the appropriate conditions OLS is unbiased.

**Illustration** In a static model,

$$y_t = \alpha_0 + \delta_0 z_t + u_t,$$

the OLS estimators are unbiased if  $E(u_t \mid \text{all } z's) = 0$ .

- If the regressors contain lagged dependent variables (e.g.,  $y_{t-1}, y_{t-1}, \dots$ ), OLS is not unbiased.

**Illustration** Consider the AR(1) model

$$y_t = \alpha y_{t-1} + u_t, (t = 2, \dots, T).$$

The OLS estimator of  $\alpha$  is

$$\hat{\alpha} = \frac{\sum_{t=2}^T y_{t-1} y_t}{\sum_{t=2}^T y_{t-1}^2},$$

which gives

$$\hat{\alpha} - \alpha = \frac{\sum_{t=2}^T y_{t-1} u_t}{\sum_{t=2}^T y_{t-1}^2}.$$

Here

$$E(\hat{\alpha}) - \alpha = E\left(\frac{\sum_{t=2}^T y_{t-1} u_t}{\sum_{t=2}^T y_{t-1}^2}\right) \neq \frac{\sum_{t=2}^T y_{t-1} E(u_t)}{\sum_{t=2}^T y_{t-1}^2}$$

because  $y$ 's are random variables related to  $u$ .

# Variances of OLS Estimators

- The same as in Chapter 3 as long as  $E(u_t u_s | \text{all } x's) = 0$  for  $t \neq s$ .
- OLS remains BLUE if
  - (i)  $E(u_t | \text{all } x's) = 0$
  - (ii)  $E(u_t u_s | \text{all } x's) = 0$  for  $t \neq s$ .If the regressors contain lagged dependent variables, the first assumption cannot hold.
- With the additional assumption of normal errors, inference is the same.



- Economic time series often have a trend.
- Two series may be trending together without any causal relation.
- Often, both will be trending because of other unobserved factors.

## Example

Number of marriages and GDP may trend together. But this is due to the third variable, population.

# Trending Time Series

- Even if those factors are unobserved, we can control for them by directly controlling for the trend.
- One possibility is a linear trend, which can be modeled as

$$y_t = \alpha_0 + \alpha_1 t + e_t, \quad t = 1, 2, \dots$$

- Another possibility is an exponential trend, which can be modeled as

$$\ln(y_t) = \alpha_0 + \alpha_1 t + e_t, \quad t = 1, 2, \dots$$

- Another possibility is a quadratic trend, which can be modeled as

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t, \quad t = 1, 2, \dots$$

# Detrending

- Adding a linear trend term to a regression is the same thing as using “detrended” series in a regression.
- Detrending a series involves regressing each variable in the model on  $t$ .
- $\hat{\beta}$  from the OLS regression

$$y_t = \hat{\alpha}_0 + \hat{\alpha}_1 t + \hat{\beta} x_t + \hat{u}_t$$

and  $\bar{\beta}$  from the OLS regression

$$\bar{y}_t = \bar{\beta} \bar{x}_t + \bar{w}_t,$$

where  $\bar{y}_t$  and  $\bar{\beta} \bar{x}_t$  are residulas from the OLS regressions

$$y_t = \hat{\gamma}_0 + \hat{\gamma}_1 t + \bar{y}_t;$$

$$x_t = \hat{\delta}_0 + \hat{\delta}_1 t + \bar{x}_t,$$

are the same.

- Time-series regressions tend to have very high  $R^2$ .
- The  $R^2$  from a regression on detrended data better reflects how well the  $x_t$ 's explain  $y_t$ .

- Often time-series data exhibits some periodicity, referred to as seasonality.

## Example

Quarterly data on retail sales will tend to jump up in the 4th quarter.

- Seasonality can be dealt with by adding a set of seasonal dummies.
- As with trends, the series can be seasonally adjusted before running the regression.
- In many applications, the method of X-12 ARIMA is used for deseasonalizing.