ECONOMETRICS

M.A. Program in Economics
YTU
Department of Economics

Time Series Data

Time Series Data

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t$$

Plan

- Differences between cross section and time series data
- Examples of time-series models
- Finite sample properties of OLS with time-series data
- Hypothesis testing
- Trends and seasonality

Time Series vs. Cross Sectional

- Time series data has a temporal ordering, unlike cross-section data
- Will need to alter some of our assumptions to take into account that we no longer have a *random sample of individuals*
- Instead, we have one realization of a stochastic (i.e. random) process

Example

Partial Listing of Data on U.S. Inflation and Unemployment Rates, 1948–1996

Year	Inflation	Unemployment
1948	8.1	3.8
1949	−1.2	5.9
1950	1.3	5.3
1951	7.9	3.3
⋮	⋮	⋮
1994	2.6	6.1
1995	2.8	5.6
1996	3.0	5.4

Nature of time series data

- How should we think about randomness in time series data?
- We can think of time series as outcomes as random variables because we do not know what the value in the next period will be? E.g. Real GDP or next month's inflation, tomorrow's closing prices at Istanbul Stock Exchange etc.
- Formally, a sequence of random variables indexed by time is called a stochastic process or time series process.
- When we collect a times series data set we obtain one possible outcome, or realization, of the stochastic process.

Examples of Time Series Models

- A static model relates contemporaneous variables:

$$y_t = \beta_0 + \beta_1 z_t + u_t$$

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

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t$$

- More generally, a finite distributed lag model of order *q* will include *q* lags of *z*

Example: Static Phillips Curve

- Phillips Curve relates unemployment to inflation:

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

- Assume: natural rate of unemployment and inflation expectations are constant.
- We can analyze contemporaneous trade-off between inflation and unemployment

7

Finite Distributed Lag Models

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + \ldots + \delta_q z_{t-q} + u_t$$

- δ_0 is called the **impact propensity** – it reflects the immediate change in *y*
- For a temporary, 1-period change, y returns to its original level in period $q+1$
- We can call $\delta_0 + \delta_1 + \ldots + \delta_q$ the **long-run propensity (LRP)** – it reflects the long-run change in *y* after a permanent change

8

Finite Distributed Lag Models

- We can additional explanatory variables into an FDL model, e.g., x(t), w(t).
- Example : Using annual data on interest rate (int) and inflation rate (inf) the following FDL model is estimated. What is the impact multiplier and LRP?

$$int_t = 1.6 + .48\, inf_t - .15\, inf_{t-1} + .32\, inf_{t-2} + u_t,$$

- Impact multiplier: 0.48
- LRP = {0.48 + (-0.15) + 0.32}= 0.65

9

Assumptions for Unbiasedness

- Still assume a model that is linear in parameters:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t$$

- Still need to make a zero conditional mean assumption:

$$E(u_t|\boldsymbol{X}) = 0,\ t = 1, 2, \ldots, n$$

- Note that this implies the error term in any given period is uncorrelated with the explanatory variables **in all time periods**

10

Assumptions (continued)

- This zero conditional mean assumption implies the x's are **strictly exogenous**
- An alternative assumption, more parallel to the cross-sectional case, is

$$E(u_t|\boldsymbol{x}_t) = 0$$

- This assumption would imply the x's are **contemporaneously exogenous**
- Or, x's are said to be **weakly exogenous**
- Contemporaneous exogeneity will only be sufficient in large samples

11

Assumptions (continued)

- Still need to assume that no *x* is constant, and that there is no perfect collinearity
- Note we have skipped the assumption of a random sample
- The key impact of the random sample assumption is that each u_i is independent
- Our strict exogeneity assumption takes care of it in this case

12

Unbiasedness of OLS

- Based on these 3 assumptions, when using time-series data, the OLS estimators are unbiased
- Thus, just as was the case with cross-section data, under the appropriate conditions OLS is unbiased
- Omitted variable bias can be analyzed in the same manner as in the cross-section case

13

Variances of OLS Estimators

- Just as in the cross-section case, we need to add an assumption of homoskedasticity in order to be able to derive variances
- Now we assume

$$\mathrm{Var}(u_t|\boldsymbol{X}) = \mathrm{Var}(u_t) = \sigma^2$$

- Thus, the error variance is independent of all the x's, and it is constant over time
- We also need the assumption of no serial correlation:

$$\mathrm{Corr}(u_t, u_s|\ \boldsymbol{X})=0 \text{ for } t \neq s$$

14

OLS Variances (continued)

- Under these 5 assumptions, the OLS variances in the time-series case are the same as in the cross-section case. Also,
- The estimator of σ^2 is the same
- OLS remains BLUE
- With the additional assumption of normal errors, inference is the same

15

Example

- Variables: **i3**, interest rate on 3-month treasury bill; **inf** : annual consumer inflation, **def** : ratio of federal budget deficit to GDP. Period: 1948-1996.

$$\begin{aligned} \widehat{i3}_t &= \underset{(0.44)}{1.25} + \underset{(.076)}{.613}\ inf_t + \underset{(.118)}{.700}\ def_t \\ n &= 49,\ R^2 = .697,\ \bar{R}^2 = .683. \end{aligned}$$

- Ceteris paribus, one percentage point change in inflation leads to a 0.613 point increase in short-term interest rates.
- We see that t-statistics are very large, therefore they are statistically significant. These can be used for inference as the classical assumptions are satisfied.

16

Example: Effects of personal tax exemption on fertility rates

- Variables: **gfr**: general fertility rate, the number of children born to every 1000 women of child-bearing age; **pe**: personal tax exemption ($); **ww2**: dummy variable for WW2, **pill**: dummy variable for birth control pill (equals 1 after 1963). Data is for the period 1913-1984.

$$gfr_t = \beta_0 + \beta_1 pe_t + \beta_2 ww2_t + \beta_3 pill_t + u_t,$$

$$\begin{aligned} \widehat{gfr}_t &= \underset{(3.21)}{98.68} + \underset{(.030)}{.083}\ pe_t - \underset{(7.46)}{24.24}\ ww2_t - \underset{(4.08)}{31.59}\ pill_t \\ n &= 72,\ R^2 = .473,\ \bar{R}^2 = .450. \end{aligned}$$

17

Example: Effects of personal exemption on fertility rates

- All coefficients are statistically significant against two-sided alternative at 1% level.
- During WW2, ceteris paribus, the number of births per 1000 women has decreased by about 24. Since the range of gfr variable is 65-127 this is a quite large impact.
- Again, ceteris paribus, introduction of birth control pills has reduced the general fertility rate.
- The mean of personal tax exemption is $100.4 and min=0 and max=$243.8. A $12 increase in PE increases the number of births per 1000 women by 1.(12 x 0.083 = 1).

18

Example: Effects of personal exemption on fertility rates

- Should we include lags of PE? GFR may respond to tax exemptions after a period of time. Estimate an FDL model:

$$\widehat{gfr}_t = \underset{(3.28)}{95.87} + \underset{(.126)}{.073}\, pe_t - \underset{(.1557)}{.0058}\, pe_{t-1} + \underset{(.126)}{.034}\, pe_{t-2}$$
$$- \underset{(10.73)}{22.13}\, ww2_t - \underset{(3.98)}{31.30}\, pill_t$$
$$n = 70,\ R^2 = .499,\ \bar{R}^2 = .459.$$

F test of excluding pe(t-1) and pe(t-2)
pvalue=0.95. Jointly insignificant
Prefer static model

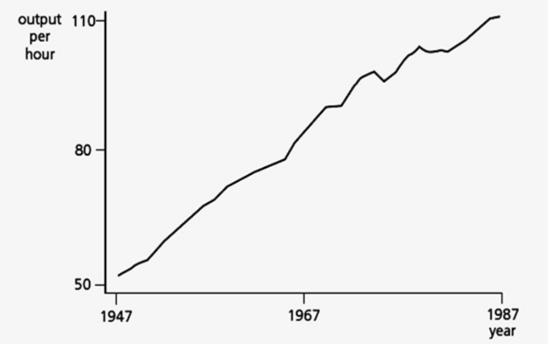
19

Trending Time Series

- Economic time series often have a trend
- Just because 2 series are trending together, we can't assume that the relation is causal
- Often, both will be trending because of other unobserved factors
- Even if those factors are unobserved, we can control for them by directly controlling for the trend

20

Output per labor hour in the United States during the years 1947–1987; 1977 = 100.



21

Trends (continued)

- One possibility is a linear trend, which can be modeled as

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

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

$$\log(y_t) = \alpha_0 + \alpha_1 t + e_t,\ t = 1, 2, \ldots$$

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

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

22

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
- The residuals form the detrended series
- Basically, the trend has been partialled out

23

Detrending (continued)

- An advantage to actually detrending the data (vs. adding a trend) involves the calculation of goodness of fit
- Time-series regressions tend to have very high R^2, as the trend is well explained
- The R^2 from a regression on detrended data better reflects how well the x_t's explain y_t

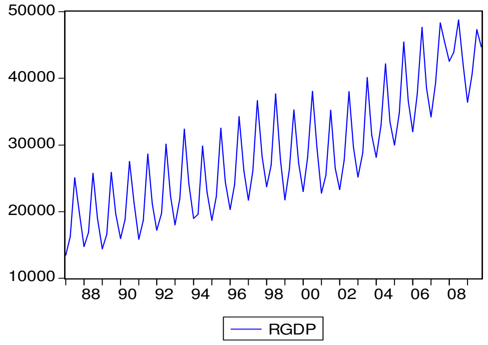
24

Seasonality

- Often time-series data, especially monthly and quarterly data, exhibits some periodicity, or cyclical movements, referred to 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

25

Example: Turkish Real GDP 1987.Q1-2009.Q4



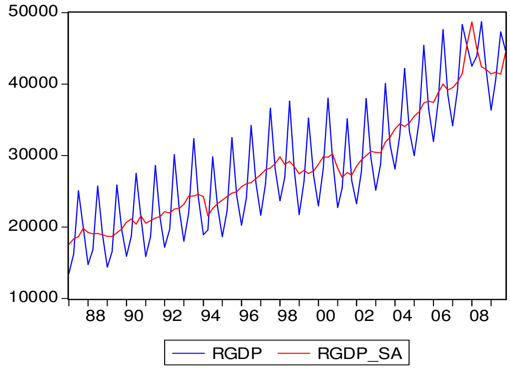
26

Example: RGDP

- Real GDP has a clear upward trend and seasonal effects.
- We can remove seasonality using either dummy variables or more sophisticated procedures in Eviews
- Eviews→PROC→SEASONAL ADJUSTMENT
- Has four SA procedures:
  - Census X12
  - X11
  - TRAMO/SEATS
  - Moving Average Methods

27

Example: Seasonally Adjusted RGDP using Census X12



28