Econometrics: Multiple Regression and Applications

ECON4004: LAB 5

Duong Trinh

University of Glasgow

March 6, 2024

Intro

- Duong Trinh
  - PhD Student in Economics (Bayesian Microeconometrics)
  - Email: Duong.Trinh@glasgow.ac.uk

- ECON4004-LB01
  - Wednesday 10am -12 pm
  - 5 sessions (7-Feb, 14-Feb, 21-Feb, 28-Feb, 6-March)
  - ST ANDREWS:357
- ECON4004-LB02
  - Wednesday 12-2 pm
  - 5 sessions (7-Feb, 14-Feb, 21-Feb, 28-Feb, 6-March)
  - ST ANDREWS:357

Record Attendance

Plan for LAB 5

- Excercise: based on Stock & Watson, E15.1
- We will focus on *"Time series Regression"*

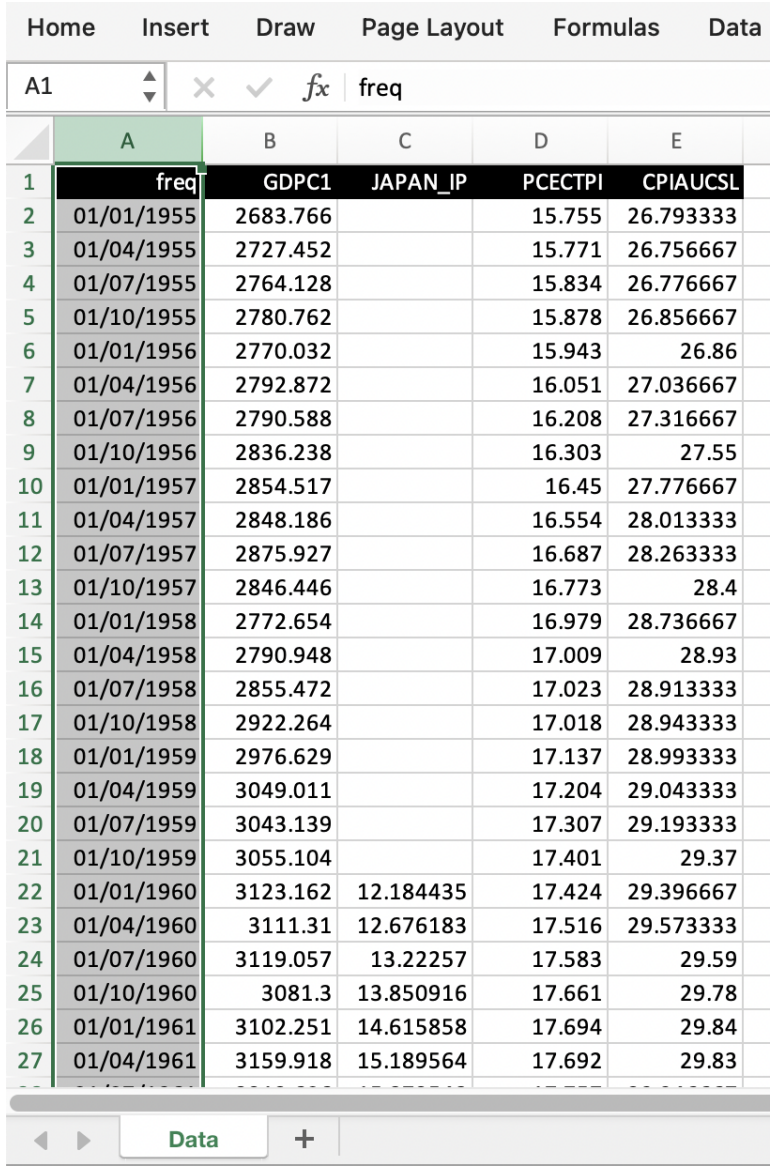
BRIEF REVIEW

Time Series Data - What it looks like...

Time series data are data collected on the same observational unit at multiple time periods (t).

$$\{Y_t\},$$
$$t = 1, \ldots, T$$

Can be of any time frequency - daily, monthly, quarterly, annual, etc.[1]



	A	B	C	D	E
1	freq	GDPC1	JAPAN_IP	PCECTPI	CPIAUCSL
2	01/01/1955	2683.766		15.755	26.793333
3	01/04/1955	2727.452		15.771	26.756667
4	01/07/1955	2764.128		15.834	26.776667
5	01/10/1955	2780.762		15.878	26.856667
6	01/01/1956	2770.032		15.943	26.86
7	01/04/1956	2792.872		16.051	27.036667
8	01/07/1956	2790.588		16.208	27.316667
9	01/10/1956	2836.238		16.303	27.55
10	01/01/1957	2854.517		16.45	27.776667
11	01/04/1957	2848.186		16.554	28.013333
12	01/07/1957	2875.927		16.687	28.263333
13	01/10/1957	2846.446		16.773	28.4
14	01/01/1958	2772.654		16.979	28.736667
15	01/04/1958	2790.948		17.009	28.93
16	01/07/1958	2855.472		17.023	28.913333
17	01/10/1958	2922.264		17.018	28.943333
18	01/01/1959	2976.629		17.137	28.993333
19	01/04/1959	3049.011		17.204	29.043333
20	01/07/1959	3043.139		17.307	29.193333
21	01/10/1959	3055.104		17.401	29.37
22	01/01/1960	3123.162	12.184435	17.424	29.396667
23	01/04/1960	3111.31	12.676183	17.516	29.573333
24	01/07/1960	3119.057	13.22257	17.583	29.59
25	01/10/1960	3081.3	13.850916	17.661	29.78
26	01/01/1961	3102.251	14.615858	17.694	29.84
27	01/04/1961	3159.918	15.189564	17.692	29.83

[1] Typical resource: https://fred.stlouisfed.org/

[SN] Working with dates and times in STATA

Date types in Stata[2]

Date type	Format	Unit
Datetime	%tc	Milliseconds since 01jan1960 00:00:00.000
Daily date	%td	Days since 01jan1960
Weekly date	%tw	Weeks since 1960w1
Monthly date	%tm	Months since 1960m1
Quarterly date	%tq	Quarters since 1960q1

[2] See guideline at: https://www.stata.com/bookstore/dtguide.pdf

[SN] Our Example

```
. * Import Quarterly Data from Excel file
. * first row considered as variable names
. import excel "us_macro_quarterly.xlsx", sheet("Data") firstrow clear
(5 vars, 252 obs)

. describe

Contains data
 Observations:           252
    Variables:             5
```

Variable name	Storage type	Display format	Value label	Variable label
freq	int	%td..		**freq**
GDPC1	double	%10.0g		**GDPC1**
JAPAN_IP	double	%10.0g		**JAPAN_IP**
PCECTPI	double	%10.0g		**PCECTPI**
CPIAUCSL	double	%10.0g		**CPIAUCSL**

[SN] Working with dates and times in STATA

(I) Option 1: Building dates and times from components.[3]

Date type	Format	Pseudofunction	Function
Daily date	%td	td(day-month-year)	mdy(M, D, Y)
Weekly date	%tw	tw(year-week)	yw(Y, W)
Monthly date	%tm	tm(year-month)	ym(Y, M)
Quarterly date	%tq	tq(year-quarter)	yq(Y, Q)

[3] See guideline at: https://www.stata.com/bookstore/dtguide.pdf

[SN] Our Example

```
. * Create a desired quarterly date, e.g. 1955q1
. display %tq tq(1955q1) //using pseudofunction tq(.)
1955q1

. display %tq yq(1955,1) //using function yq(.)
1955q1

.
. * Generate quarterly date variables, recursively starting from 1955q1
. * By default, tq(1960q1) is defined to be 0 in Stata.
. gen date1 = tq(1955q1) + _n-1

. gen date2 = yq(1955,1) + _n-1

. format %tq date1 date2 // express them in quarterly format, see Data Editor

. list if date1 != date2 // check if both variables are identical
```

[SN] Working with dates and times in STATA

(II) Option 2: Converting dates and times from existing variables.[4]

		To		
From	Daily date	Weekly date	Monthly date	Quarterly date
Daily date		`wofd()`	`mofd()`	`qofd()`
Weekly date	`dofw()`		`mofd(dofw())`	`qofd(dofw())`
Monthly date	`dofm()`	`wofd(dofm())`		`qofd(dofm())`
Quarterly date	`dofq()`	`wofd(dofq())`	`mofd(dofq())`	

[4] See guideline at: https://www.stata.com/bookstore/dtguide.pdf

[SN] Our Example

[SN] STATA command for Setting Data as Time Series

Excercise: based on Stock & Watson, E15.1

Picture the Scenario

- **Objective:** Construct forecasting models for the rate of inflation.
- **Dataset:** `us_macro_quaterly.xlsx`.
  - contains quarterly data on several macroeconomic series for the US.
  - use the sample period $1963 : Q1 - 2017 : Q4$.(where data before 1963 may be used, as necessary, as initial values for lags in regressions)
- **Key variables:** For each country in each year
  - `PCEPI`: the price index for personal consumption expenditures from the U.S. National Income and Product Accounts.

Questions

(a)

i. Compute the inflation rate, $Infl = 400 \times [ln(PCEPI_t) - ln(PCEPI_{t-1})]$. What are the units of `Infl`?

ii. Plot the values of `Infl` from 1963:Q1 through 2017:Q4. Based on the plot, do you think that `Infl` has a stochastic trend? Explain.

(b)

i. Compute the first four autocorrelations of $\Delta Infl$.

ii. Plot the value of $\Delta Infl$ from 1963:Q1 through 2017:Q4. The plot should look choppy or jagged. Explain why this behaviour is consistent with the first autocorrelation that you computed in (i.).

Questions

(c)

i. Run an OLS regression of $\Delta Infl_t$ on $\Delta Infl_{t-1}$. Does knowing the change in inflation over the current quarter help predict the change in inflation over the next quarter? Explain.

ii. Estimate an $AR(2)$ model of $\Delta Infl_t$. Is the $AR(2)$ model better than the $AR(1)$ model? Explain.

iii. Use the $AR(2)$ model to predict the change in inflation from 2017:Q4 to 2018:Q1 - that is, to predict the value of $\Delta Infl_{2018:Q1}$.

Questions

(d)

i. Use the ADF test for $AR(p)$ regression

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \ldots + \gamma_p \Delta Y_{t-p+1} + u_t$$

using 2 lags of $\Delta Infl$ (so that $p = 3$ in the above equation) to test for a stochastic trend in *Infl*.

ii. Is that ADF test based on the above regression preferred to the regression including a determinstic trend

$$\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \ldots + \gamma_p \Delta Y_{t-p+1} + u_t$$

for testing for a stochastic trend in *Infl*? Explain.

iii. Based on the ADF tests carried out, does the AR model for *Infl* contain a unit root?