ECON 202 Intermediate Macroeconomics: Basic Tools of Empirical Macroeconomics Lec 9-12

Sonan Memon
Lecturer, Institute of Business Administration, Karachi.

February 2021

What is Time Series Data?

i. There are no explicitly required set of readings for this lecture but you may find reading the first few pages of chapter 20 from GLS (2020) helpful.

ii. Time series data $\{x_t\}_{t=1}^{T}$ consists of just a series of data observations indexed by time.

iii. Order of time series data has meaning.

iv. The observations are not likely to be independent over time due to history dependence and autocorrelation.

Introduction to Empirical Macroeconomics

i. Empirical macroeconomics refers to the set of tools and methodologies used to describe, analyze, organize and carry statistical inferences on macroeconomic data.

ii. Macroeconomic data also includes micro data such as household data or survey data, especially due to micro foundations in modern macroeconomics.

Some Workhorse Models in Empirical Macroeconomics

i. The list of available *time series*, *panel data* and *cross-sectional data* econometric methods is huge.

ii. Structural versus reduced form methods.

iii. Some workhorse models include simple ARDL models, Vector Autoregressive Models, Factor Models, small or large scale structural econometric models including GMM and SMM, ARCH and GARCH models, machine learning methods, regressions with panel data in macroeconomics, microeconometric methods applied to micro data such as survey data on expectations or firm/household data etc etc etc.

Eye Ball Econometrics

i. Of course, with time series data, the first thing we should do is to plot the time series data and eye ball it.
ii. This can give you a sense of basic patterns and facts in the data.
iii. This is not a substitute for formal statistical methods but can motivate further empirical inspections.
iv. Eye balling data may be a source of new questions, puzzles and hypotheses.

R Application

i. `ts` function.
ii. `as.Date` function.
iii. `ggplot` functionality.
iv. Some required R packages are: `tseries`, `ggplot2`, `forecast`, `dynlm`, `mFilter`, `caTools` etc.

Trends and Stationarity of Time Series

i. The first prominent thing to note about time series data is the existence of any trend over time.
ii. What kind of trend is it? Upward or downward and is the trend linear or nonlinear?
iii. The second thing to worry about is whether the time series is stationary or not.
iv. A stationary or technically speaking weakly stationary time series has constant mean and variance over time.
v. There are rigorous statistical tests that one can perform on time series to detect non-stationarity such as Augmented Dickey-Fuller test, among others.

Lead and Lag Variables

i. Lead versus lag of series: $L(x_t) = x_{t-1}$, $L^{-1}(x_t) = x_{t+1}$, L is called the Lag operator. $L^j(x_t) = x_{t-j}$ and $L^{-j}(x_t) = x_{t+j}$.

ii. First difference of series is current value minus the first lag $L(x_t) = x_{t-1}$, i.e $x_t - x_{t-1} = Diff^1(x_t)$, where $Diff^j$ is the difference operator.

iii. First difference of ln of a series x_t is $ln(x_t) - ln(x_{t-1})$, which approximates the growth rate of original series $g_{x,t}$ at time t.

Basics of Time Series Data

Time Series Data			
Date	$Murders_t$	$L^1(Murders_t)$	$Diff^1(Murders_t)$
1990	5000	NA	NA
1991	5500	5000	500
1992	6000	5500	500
1993	5700	6000	-300
...	...	...	...

R application

i. A simple plot can sometimes make the stationarity problem obvious but you should always perform formal, statistical tests as well.

ii. `adf.test` function for Augmented Dickey-Fuller Test for stationarity.

iii. p value of less than 0.05 reflects that time series is stationary.

iv. `diff` function to take difference of series. First difference of log of a series approximates the growth rate of the original series.

Moving Average of Growth Rates

Example with window $h = 3$			
Date	$g_{x,t}$	$BMAG_t^h$	$CMAG_t^h$
2004Q1	2.5	NA	NA
2004Q2	1.3	NA	$\frac{2.5+1.3+1.9}{3}$
2004Q3	1.9	$\frac{1.9+1.3+2.5}{3}$	$\frac{1.3+1.9+2.4}{3}$
2004Q4	2.4	$\frac{2.4+1.9+1.3}{3}$	$\frac{1.9+2.4+1.7}{3}$
2005Q1	1.7	$\frac{1.7+2.4+1.9}{3}$	NA

Rolling Window Growth Rates

i. If growth rates are averaged in some window h, we get moving/rolling average of growth rates over time due to rolling/changing and overlapping windows.

ii. Moving Average of Growth Rates: $BMAG_t^h = \frac{1}{h}\sum_{k=0}^{h-1} g_{t-k}$ (one sided, backward with horizon h).

iii. Centered Moving Average: $CMAG_t^h = \frac{1}{h}\sum_{k=-\frac{h-1}{2}}^{\frac{h-1}{2}} g_{t-k}$ (two sided, centered moving average with horizon h).

iv. Forward or left aligned Moving Average: $FMAG_t^h = \frac{1}{h}\sum_{k=0}^{h-1} g_{t+k}$ (one sided, forward with horizon h).

v. Why do we take moving averages?

R application

i. `runmean` function from `caTools` package.

ii. Specify horizon k in function, treatment of end points, whether the moving average is left aligned (forward), right aligned (backward) or centered and algorithm type.

Annualized Growth Rates

i. In order to make growth rates comparable, we often annualize them by converting growth rates for weekly, daily, monthly or any other frequency into annual terms.

ii. Another way to think about this is that annualized growth rate adjusts a growth rate to reflect the amount a variable would have changed over a year's time had it continued to grow at the given rate.

iii. Formula: $g_m = \left(\left(\frac{Y_m}{Y_{m-1}} \right)^{12} - 1 \right)$ where Y_m is value of economic variable in month m and g_m is the annualized growth rate from month $m-1$ to m.

Dynamic Correlations or Lead Lag Structures

i. Correlogram: quantifies the correlation between any two time series $\{x_t\}_{t=1}^{T}$ and $\{y_t\}_{t=1}^{T}$ at various leads and lags.

ii. For example $cor\ (x_{t+h}, y_t)$can be measured for various values of $h = ..., -2, -1, 0, 1, 2....$

iii. This gives us a sense of lead lag structure between any two series. Leading or lagging indicators can be identified.

iv. Of course correlation is not causation but it is a starting point for empirical analysis of causation. If you do not even have correlation, then causation cannot occur.

R application

i. `cor` function for simple correlation.

ii. `ccf` function for cross-correlations of form $cor\left(x_{t+h}, y_t\right)$ for horizons $h = ..., -2, -1, 0, 1, 2....$

iii. Plot cross-correlation function, also sometimes called dynamic correlations or the lead lag structure between any two series.

Quantifying Volatility and Relative Volatility of Time Series Data

i. Standard deviation of time series over some sample is used as a simple method for quantifying volatility of a time series: $\sigma_x = SD\left(\{x_t\}_{t=1}^{T}\right)$.

ii. $\sigma_x = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(x_t - \bar{x}\right)^2}$, $\bar{x}$ is mean.

iii. Volatility is often measured relative to volatility of some other benchmark such as output growth rate y: $\frac{\sigma_x}{\sigma_y}$.

iv. Volatility can also be time varying: $\sigma_{x,t}$. Time varying and stochastic volatility models are very complicated for this level. ARCH/GARCH models are examples.

v. Rolling standard deviations serve as quick and dirty methods for understanding time varying volatility.

Rolling Standard Deviations

i. One sided measure of volatility with horizon h: $BMVG_t^h = SD\left(\{g_{t-k}\}_{k=0}^{h-1}\right)$ (one sided, backward with horizon h and center t).

ii. Centered Measure of Volatility: $CMVG_t^h = SD\left(\{g_{t-k}\}_{k=-\frac{h-1}{2}}^{\frac{h-1}{2}}\right)$ (two sided, centered with horizon h and center t).

iii. Forward or left aligned Moving Volatility: $FMVG_t^h = SD\left(\{g_{t+k}\}_{k=0}^{h-1}\right)$ (one sided, forward with horizon h and center t).

iv. Why define this concept?

R application

i. `runsd` function from `caTools` package.

ii. specify data series, center, horizon, alignment and other usual suspects as function arguments.

Trend Versus Cycle Separation

i. Trend versus cycle separation: measuring business cycles as deviations from trend. Linear trends?
ii. Hodrick Prescott filter provides one method for trend, cycle separation.
iii. Many other methods such as Band pass filter and Baxter-King filter etc which we will not cover.
iv. The idea is to separate out variation at both too high and too low frequencies, let's say more than 32 quarters or less than 6 quarters to isolate meaningful business cycle fluctuations.

Linear Trend Separation

i. $x_t = \tau_t + \tilde{x}_t$. where x_t is time series, τ_t is trend component and $\tilde{x}_t$ is cyclical component which is $x_t - \tau_t$.

ii. In order to extract $\tilde{x}_t$, we need to identify τ_t since data on x_t is already available.

iii. Linear trend $\implies$ $\tau_t = \widehat{a} + \widehat{b} \times t$ where t is just time periods and can be estimated using simple, linear regression.

HP Filter

i. In order to extract $\tilde{x}_t$, we need to identify τ_t since data on x_t is already available.

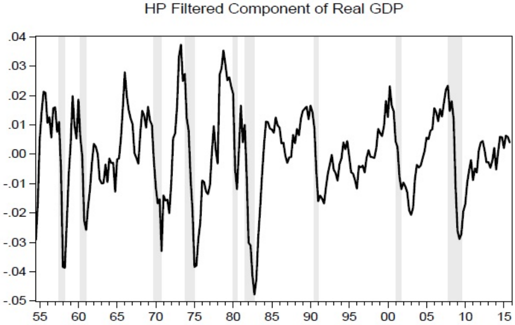
ii. HP filter gives another method for choosing τ_t by solving the following optimization problem:

iii. $\underset{\{\tau_t\}}{min} \sum_{t=1}^{T} (x_t - \tau_t)^2 +$
$\lambda \left(\sum_{t=2}^{T-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}) \right]^2 \right)$.

iv. HP filter is governed by smoothing parameter λ and $\lambda = 1600\ (6.25)$ should be used for quarterly (annual) time series. $\lambda = 129600$ for monthly series.

HP Filtered US GDP

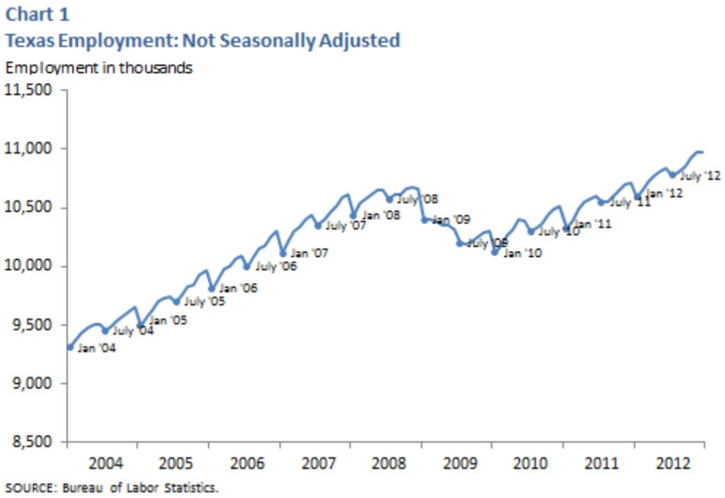
Figure 20.1: Cyclical Component of Real GDP



Seasonality

i. Studying economic variables over time helps show the extent to which a particular industry of the economy is contributing to economic growth for example.
ii. 1 problem with interpreting data over time is that many data exhibit movements that recur every year in the same month or quarter.
iii. For example, housing permits increase every spring when the weather improves, while toy sales usually peak in holidays.
iv. This makes it hard for economists to interpret the underlying, **fundamental** trend in some data series. Example: Eid Holidays.

Seasonality in Texas Employment Data



Seasonal Adjustments

i. Economists want to know if sales were **better than the normal seasonal increase** $\implies$ remove predictable fluctuations/seasonality from data.

ii. One way to do this is to compare monthly or quarterly data on a year-over-year basis. In other words, the current month's data point is compared with the data point from the same month in the prior year.

iii. However, year over year growth rate relies heavily on data that is 12 months old but comparison from month to month helps economists determine significant changes in the business cycle soon after they occur.

Seasonal Adjustments

i. Hence, we most commonly use the $X12$ procedure, developed by the US Census Bureau to adjust data for seasonal patterns.

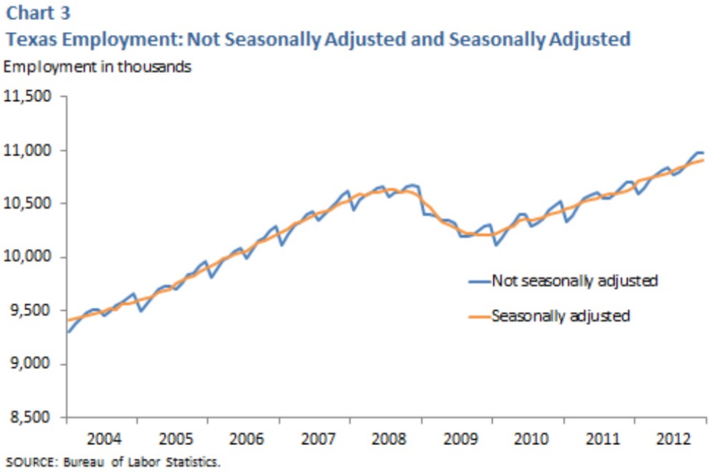
ii. $x_t = \tau_t + \tilde{x}_t + irregular_t + seasonal_t$

iii. The trend-cycle component is the series' long-term tendency to grow or decline and can fluctuate because of the economic trends or other long-term cyclical factors.

iv. The irregular component comes from unseasonable weather, natural disasters, strikes or sampling error.

v. The goal of the seasonal adjustment procedure is to separate out the seasonal component, leaving the trend-cycle and irregular components.

Seasonally Adjusted Texas Employment Data



R application

i. `lm` function and `resid` function for linear detrending.
ii. `hpfilter` function from `mFilter` package for HP filtering.

Thank you