Introductory Econometrics

Tutorial 2

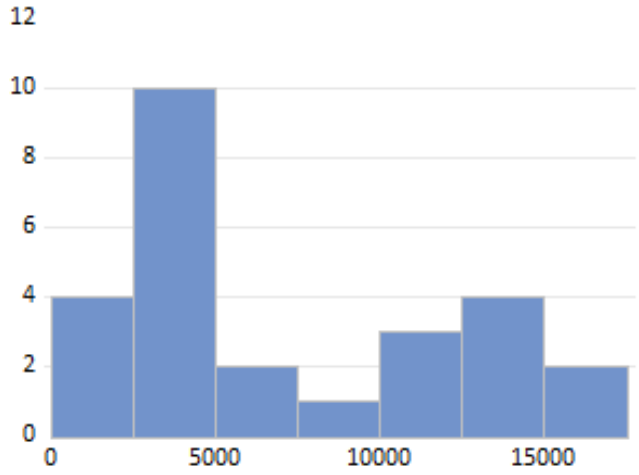
PART A:

This part was mostly an introduction to a data set that you will be using for the first assignment. The results for the application of VLOOKUP function was already in the hidden sheet of the data file.

PART B:

Question 1: *Important topics of the first year statistics: Histograms, scatter plots, and remembering logarithms*: Read the data that you compiled for "A sample of countries" into EViews. When an Excel spreadsheet contains several sheets of data, the first dialogue in EViews allows you to specify which sheet you are interested to read in EViews. If you have not saved your Excel work or had difficulty with it, download "A sample of countries.wf1" from the moodle site and use that.

1. Obtain the summary statistic and histogram of GDPPC. Discuss what you can learn from the histogram and summary statistics (remember that this is not a representative sample of all countries in the world, but is a random sample of countries with GDP per capita of less than 50 dollars a day). If you had a different set of 26 countries with GDP per capita of less than 50 dollars a day, would the summary statistics be the same?

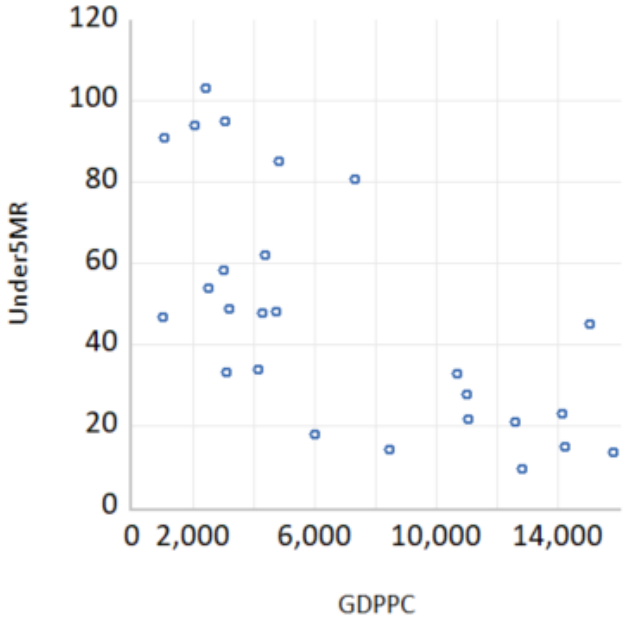


Series: GDPPC	
Sample 1 26	
Observations 26	
Mean	7031.317
Median	4785.331
Maximum	15788.82
Minimum	1037.546
Std. Dev.	4847.148
Skewness	0.492281
Kurtosis	1.741225
Jarque-Bera	2.766700
Probability	0.250737

The histogram is bi-modal. Unfortunately the larger mode is below 5000. Even in this group of countries, there is a large disparity of income, with more than half of the countries in the sample having a GDP per capita of less than 5000 dollars (the median shows that half of the countries had GDP per capita less than or equal to 4785.331 dollars. The sample average here is an estimate of the mean income for the population of countries with GDP per capita of less than 50 dollars a day. If we had a different set of countries, even from the list of countries with GDP per capita of less than 50 dollars a day, we will end up with a different set of summary statistics. These statistics are estimates of population parameters, and while the population parameters are fixed, these estimates change as the sample changes (hence they are random variables).

2. We want to explore the association between under_5 mortality and GDP per capita for lower than median income countries. What kind of a graph can give us an insight into the nature of this relationship? Based on this graph, are under_5 mortality and GDP per capita positively

or negatively correlated? Is their relationship linear?



A scatter plot can reveal the nature of correlation between two variables. The above shows that under-5 mortality rate and GDP per capita are negatively correlated (have an inverse relationship). It is not that easy to determine linearity or nonlinearity based on 26 observations. However, in this case, thinking about the problem will give us a clue. Under-5 mortality has a lower bound of zero. It can never be negative (and achieving 0 is also impossible because we cannot avoid all possible accidents). So, we know that a downward sloping linear model will not be appropriate. We will study and think about functional forms later in this unit.

Question 2: *Working with time series: plots, trends, seasonality, growth rate (log-returns):* Download hourly wages from www.ausmacrodata.org-> Categories -> wage price index -> the first series that shows up -> Download CSV. Open the CSV file, and tidy up the data set, i.e. only keep the first two columns and delete everything else, and give a better name for the second column, such as WAGE. Save the CSV file and read it in EViews. Note that EViews automatically realises that you have quarterly data.
About ausmacrodata.org: This website was created by me and two other staff members and an honours student, with the financial support of the Australian Research Council. It automatically updates itself at 3 am every day by crawling over the ABS and Reserve Bank data websites. Students may like this story because it tells them how close they are to being able to produce something that can be used by all researchers who want to work on Australian macro economy.

1. Plot the WAGE series (plotting a time series means producing a line plot of the series in which the x-axis is time. In EViews, clicking on a series opens a window that shows the value of the series in a spreadsheet. This window has a menu bar. Under View is Graph. And the default for Graph is a line plot). What can we learn from this plot?

The plot shows that hourly wage has been growing steadily in this period (positive or upward trend). Explain this means high persistence (each observation very close to observation before it). Although not so clear, some may see accelerating growth in the first half and decelerating growth towards the end.

2. If we were interested in forecasting hourly wage in the next period, would sample average be a good forecast? Suggest more appropriate forecasts.

   No, sample mean will not be of much use here. The population mean is changing over time and increasing with time. We can use the last observation as the forecast for the next period, but we can do better than that. For example, we can compute the average quarter on quarter growth of wage over the sample period, and assumes that wages grow by that amount next quarter and make a better forecast based on the end quarter and this growth rate. We can even do better than that given what we learn in part 3.

3. In most financial or economic time series, trend is so dominant that it is the only thing that we can immediately see and all other aspects of the time series are dwarfed by its trend. To see other aspects of the series we have to remove its trend. One way is to make a model with time as an explanatory variable (we will do this later in the course). Another way is to compute the growth rate, in this case $g_t = 100 \times \frac{wage_t - wage_{t-1}}{wage_{t-1}}$ (multiplication by 100 is just to express it in percentage points). A more prevalent way of calculating the growth rate, in particular in finance, is to use what is known as "log-returns"

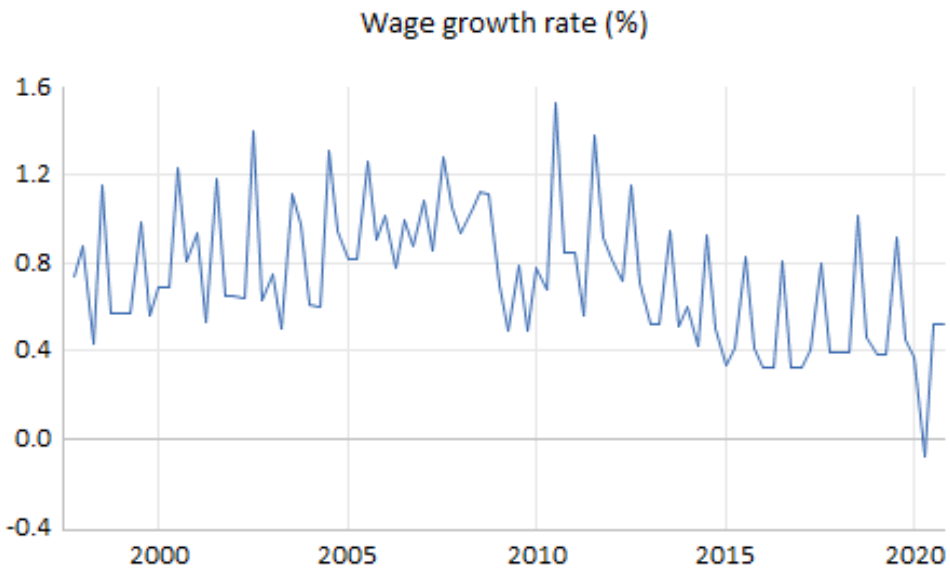
$$g_t = 100 \times \Delta \log(wage_t) = 100 \times (\log(wage_t) - \log(wage_{t-1}))$$

   These two methods of calculating the growth rate produce values that are close to each other as long as growth rate is less than 10% in absolute value. EViews has a built in function 'dlog(X)' that computes the difference of logarithm of X. Generate the growth rate of hourly wage using the log-returns formula. Open this series. Why is the first value of this series NA?

	Modified: 9/01/1997 3/01/2018 // g=100*dlog(wage)
9/01/1997	NA
12/01/1997	0.739102
3/01/1998	0.879771
6/01/1998	0.437000
9/01/1998	1.156082

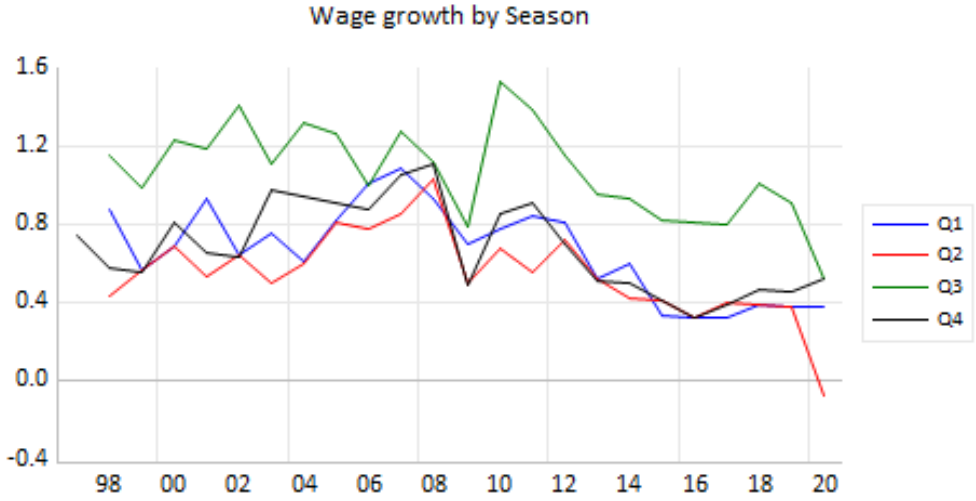
Because to compute $100 \times (\log(wage_t) - \log(wage_{t-1}))$ when t is the beginning of the sample (1997 Q3), we need wage data for the previous period (1997 Q2), which we do not have.

4. Plot the growth rate of wage. What does this plot tell you?



There is a clear seasonality (intra-year pattern) in the data. If they hover the mouse over the peaks, they can see that most (if not all) correspond to the 3rd quarter of each year. This may be due to the fact that most people's wage increases become effective at the beginning of the financial year. This is my guess. Also note the COVID effect.

5. Look at the seasonal plots of the growth rate of wage. To produce seasonal plots in EViews, in the series window, View -> Graph, and choose Seasonal Graph (the last option under Graph type). There are two seasonal plots: one that plots each season in a different panel side by side and shows average growth rate for each season. Another type shows four line plots, one for each season, overlayed on one graph. Look at both plots and discuss what you learn from these plots.



Both plots show that wage growth in Q3 is higher than in other quarters.

6. Look at the time series plot of the growth rate of wage again. Mentally adjust for seasonal variation. Do you see that the wage growth has been declining since 2010? Should that by itself worry us? What else do we need if we are worried about the value of one hour of work?

   The point here is that these are nominal wages. The real value of work depends on how much we can buy with our wages. If nominal wage growth is lower but inflation is lower as well, the real wage can be lower, steady or higher depending on the relative magnitudes of nominal growth rate and inflation rate. So, we need to look at the inflation rate to determine what has happened to the real wage growth.