Econ 303 Project 1 -- Macroeconomic Data Exploration 1/19/2009

Wei-Choun Yu

1. Due **Feb 6** (F) 6:00pm. No late submission will be accepted.
2. Total scores: 100 points. You can discuss the project with your classmates *but you need to do the project on your own hands and write down your own answers*. Otherwise, you learn little form this project. If I find the exact same answers you turn in, you will get penalty on project's grades.
3. In this project, you will learn how to use MS EXCEL spreadsheet to master the first-hand macroeconomic data and interpret them. When you have no clues how to explain the questions, check the textbook. For learning advanced skills on Excel, you can find a lot of books about Excel in our library. Also you will find a bunch of useful materials (such as video tutorials) in the following links: Atomic Learning, DataPig Excel Training and Excel Internet Library. However, most of them are based on Excel 2003 edition. My guideline is based on **Excel 2007 edition**. Be patient about using Excel. In this project, you might encounter a lot of trouble. It happens to everyone. If that happens, try again and use different ways. If you have tried 30 minutes and it still doesn't work, ask your classmate or me for help.
4. Go to the course website to download the Excel file **project1.xls** (next to this guideline) to your personal disk. Open the file and type your name in Cell C1 of "Essay" Tab.
5. After finishing your homework, turn in your project **electronically** via D2L.
   (1) From the *Course Home* page, on the *Navigation* bar, click **DROPBOX**
   The *Folder List* dialog box appears.
   (2) Click "Project 1" folder. The *Drop Off Files* dialog box appears.
   (3) In the *File* text box, select the file: (a) Click **BROWSE...** The *Choose file* dialog box appears. (b) Using the *Look in* pull-down list, locate and select the file. (c) Click **OPEN.**
6. Double check. View all the files that have been submitted to me in the your *Dropbox*.
   From the *Course Home* page, on the *Navigation* bar, click **DROPBOX**
   The *Folder List* dialog box appears. Click **HISTORY.** The *Dropbox Dropoff History* dialog box appears. A list of submitted files appears in each folder.
7. Finally you will have to submit one Excel file (project1.xls) with the following tabs:

- Essay – 25 points
- GDP tab -- 10 points
- Chart1 (US Real GDP Time Series) -- 5 points
- Chart2 (US Real GDP Growth Rate) -- 5 points
- Cgart3 (US Real GDP Components in '07 and '67) -- 5 points
- Chart4 (US Consumption Components) -- 5 points
- Chart5 (US Investment Components) -- 5 points
- CPI tab -- 5 points
- Chart6 (Headline and Core CPI Inflation Rate) -- 5 points
- Deflator tab -- 5 points
- Chart7 (GDP Deflator Inflation Rate) -- 5 points
- World tab -- 10 points

- Budget tab -- 5 points
- Chart8 (US Government Budget) -- 5 points

I. An Exploration of the U.S. GDP

❖ Import the Data

1. Go to the course website
2. Click Bureau of Economic Analysis under Data Resources
3. Under National section, click Gross Domestic Product (GDP) and then click Interactive Tables: National Income and Product Accounts Tables
4. Click Frequently Requested NIPA Tables
5. Click Table 1.1.6 Real Gross Domestic Product, Chained Dollars (A) (Q)
6. Choose period from "1948-A&Q" to "2007-A&Q" and check "**Quarterly (Q)"** and click "Update."
7. Click "Download Dates Shown (CSV)" →Click "Save" when the small window pops up.
8. Save the NIPATable.csv in your preferred folder in your computer.
9. Open the NIPATable.csv and copy the data (from row 1 to row 31) and paste them onto the tab "GDP" of project1.xls file (from Cell A1).
10. You will see the GDP and its components data from 1948-I to 2007 - IV (from Column C to Column IH and from Row 7 to Row 31).
11. Click the whole row 8 and right click mouse and then click "Format Cells…" and then choose "Patterns" and then pick up the color you like for C (consumption). Do the same thing for Investment (row 12), Government spending (row 26), Export (row 20), Import (row 23), and Net Export (row 19).

❖ Follow the Guidelines *(Based on MS Excel 2007 Edition)* Below

A. ***Chart 1:*** Plot the U.S. RGDP postwar time series

1. Highlight from **B6** through IH7 (two rows). Click "Insert" on the toolbar. And click Line on the toolbar .
2. Click the first sub chart time under 2-D Line.
3. You could click "**Chart Layouts**" on the toolbar and then you will be able to add/change the axis/legend title.
4. Name the chart title as "US Real GDP", x axis "Year" and y axis "Billion Dollar".
5. On the top right corner of the toolbar, you can see "Move Chart". [Or you can put the cursor on the chart, right click your mouse.] Choose "Move Chart" and choose New sheet: to save the chart in a new sheet.
6. To get a better score for this project, you should try to make your charts more professional. For example, see http://www.datapigtechnologies.com/flashfiles/formatcharts.html

B. ***Chart 2:*** Calculate and Plot the U.S. RGDP growth rate.

1. Go back to "GDP" tab.
2. Select cell **D40** and type "**= 400* (D7-C7) / C7"** and than hit "Enter" on the keyboard. This means GDP Growth Rate $_{1948Q2} = (GDP_{1948Q2} - GDP_{1948Q1}) / GDP_{1948Q1}$
3. *Copy the formula.* Click cell **D40** again. Move the cursor when the black cross "**+**" cursor showing up, drag the cursor **+** through **E40**, **F40**, …. to **IH40**.
4. You are supposed to see numbers on **row 40** representing US GDP growth rate.
5. *Plot the Chart 2.* Click and hold "Ctrl" on the keyboard and highlight row 6 and row 40. And the click "Insert" on the toolbar. And then do the same procedure like you did for chart 1. Name the chart title as "US Postwar Real GDP Growth Rate", x axis "Year" and y axis "Growth Rate".
6. *Compare your results to BEA's.* Go to Frequently Requested NIPA Tables in BEA. And click Table.1.1.1. Percent Change From Preceding Period in Real Gross Domestic Product (A) (Q). Again, choose period from "**1948**-A&Q" to "**2007**-A&Q" and check "**Quarterly (Q)"** and click "Update." Click Download Dates Shown (CSV). Save the file as NIPATable02.csv in your preferred folder in your computer. Open the file and choose cells from C7 to **IG** 7 and copy and paste it in the "GDP" tab file row 41 (from cell D41 to IH41) of project1.xls.
   Note: The numbers in row 40 and row 41 should be very similar. If not, it means that you got it wrong!

C. Calculate the mean and standard deviation (SD) of the real GDP growth for two periods: 1960-I to 1984-IV and 1985-I to 2007-IV.

1. *Calculate GDP Mean from 1960-1984.* From "GDP" tab of project1.xls, in cell C43, type "=Average(BC40: ET40)" and hit "Enter".
2. *Calculate GDP Mean from 1985-2007.* In cell C44, type "=Average(EU40: IH40)" and hit "Enter".
3. *Calculate GDP Standard Deviation from 1960-1984.* In cell C45, type "=Stdev(BC40: ET40)" and hit "Enter".
4. *Calculate GDP Standard Deviation from 1985-2007.* In cell C46, type "=Stdev(EU40: IH40)" and hit "Enter".

→ *Note: For more details about standard deviation, see Note 1 on page 7.*

D. The composition of GDP.

1. In Cell II8, II12, II19, II20, II23, and II26, type in "Consumption", "Investment","Net Export","Export","Import","Gov't Spending".
2. *GDP composition 2007.* Click cell IJ7 and type "**=IH7/IH$7**" and hit "Enter".
3. *Copy the formula.* Click the cell IJ7 again and move the cursor until black cross + showing up. Scroll down the cursor from IJ7 to IJ31.
4. *GDP composition 1967.* Do the same thing for 1967-II.

(Hint: In cell IK7, type "= CB7 / CB$7", and you will only see the results in cell **IK8**, IK12, IK20, IK23, and IK26; delete "#VALUE!" in other cells)
Note: If you are interested in knowing the difference between IH7 *and* IH$7*, you can read Note 4 on page 8.*

5. *Chart of Pie.* Holding "Ctrl" and highlight main components of 2007 GDP (Cell II8 and IJ8, II12 and IJ12, II20 and IJ20, II23 and IJ23, II26 and IJ26). And then click "Insert." And then click "Pie". Name this chart "2007 GDP." Put the cursor on the chart and right click the mouse, choose "Move Chart.." and then check New sheet and then type "**Chart3**" in the box and then click OK. A small question box will pop up saying "This sheet already exists…" and then click Yes. Arrange the chart to the left half of the white board.
   Go back to the GDP tab and do the same thing to 1967 GDP (highlight Cell II8 and IK8, II12 and IK12, II20 and IK20, II23 and IK23, II26 and IK26). ..click "Insert," choose "Pie"… Name the chart "1967 GDP"…→ right click and choose Move chart → Check New sheet: and type "**Chart3**" in the box and then click OK. Arrange the chart to the right half of the white board. Make 2007 and 1967 pie charts comparable.
6. Go back to the GDP tab, highlight HG6 to IH6 (2001 Q1 through 2007 Q4) and push "Ctrl" and highlight HG9 to IH9 for Durable Goods; HG10 to IH10 for Nondurable Goods; HG11 to IH 11 for Services together. And then click "Insert" and click "Column". And choose the chart sub-type. Click Next. Change the appropriate name (durable, nondurable, services) for series1, 2, and 3. Name the Chart title as "US Consumption Components". And save it as a new sheet **Chart4.**
   To change name for series 1, 2, 3, you can choose "Select Data" on the toolbar. A window pops up, and then choose Edit and type a new name.
7. Do the similar thing like you did in step 6 (Chart4) for components of Investment, which are Nonresidential (HG14-IH14), Residential (HG17-IH17) and Inventory change (HG18-IH18) from 2001 Q1 to 2007 Q4. Name the Chart "US Investment Components" and save it as **Chart5.**

E. *Essay Questions.* Answer and explain the following questions using MS Word file.

1. What does the U.S. postwar Real GDP trend look like from Chart 1? What kind of implication you can get from this trend?
2. For Chart 2, why do we use formula "400*(D7-C7) / C7" instead of "100* (D7-C7) / C7". Moreover, what implications can you get from Chart 2?
3. In Step C, we get U.S. RGDP growth mean and standard deviation for 1960-1984, and 1985-2007, respectively. What implications can you find from these numbers?
4. From Step D. 1-3 and Chart3, comparing 2007-IV with 1967-II, what implications can you get for the change of the compositions of US GDP?
5. Comparing Chart4 with Chart5, what implications can you find?

II. An Exploration of the U.S. Inflation

❖ Import the Data

1. Go to the course website and click Federal Reserve Economic Data (FRED) under Data Resources and click Consumer Price Indexes (CPI) under Categories
2. Click and download the following Excel files into your preferred folder in your computer
   - CPIAUCSL (**Headline CPI**: CPI for all urban consumers: all items; SA - seasonal adjusted).
     Note: In the CPIAUCSL website, click Download Data and click and download CPIAUCSL.xls.
   - CPILFESL (**Core CPI:** CPI for all urban consumers: all items less food & energy SA)
3. Go back to FRED website and click Gross Domestic Product (GDP) and Components
4. Click Price Indexes & Deflators under Categories
5. Click and download the following files into your preferred folder in your computer
   - GDPCTPI (Gross Domestic Product: Chain-type Price Index)

❖ Follow the Guidelines Below

A. Calculate and plot the inflation rate using monthly Headline CPI and Core CPI (CPI without food and energy).

1. Open the CPIAUCSL and CPILFESL Excel files.
2. Copy Headline CPI data (from 1958-1-1 to 2007-12-1 in Column B) from CPIAUCSL.xls and paste it onto column B (from cell B3 to B602) in "CPI" tab of project1.xls.
3. Copy Core CPI data (from 1958-1-1 to 2007-12-1 in Column B) from CPILFESL.xls and paste it onto column C (from cell C3 to C602) in "CPI" tab of project1.xls.
4. Use the same way you calculated GDP growth rate to calculate inflation rate for Headline CPI and Core CPI.
5. To calculate Headline CPI inflation, for example, click cell D4 and type "=**1200***(B4-B3)/B3" and hit Enter, and copy this formula through cell D5 to D602 by scroll down the black cross cursor.
6. Using the same method to calculate Core CPI in column E.
7. Highlight column A (A4-A602), column D (D4-D602), column E (E4-E602) together by holding "Ctrl".
8. Plot these two inflation series in a new sheet as **Chart6**. (Use the same way you plot for GDP growth rate)

B. Calculate and plot the inflation rate using GDP deflator.

1. Open the **GDPCTPI**.xls and copy the GDP deflator **data** onto Column B in "Deflator" tab of project1.xls and calculate the inflation from 1948-II to 2007-IV in Column C.
2. Plot the Inflation Rate from 1948-II to 2007-IV and save it as a new sheet named **Chart7.**
3. Similar to **Part I. C**, calculate US inflation rate mean and standard deviation for two periods in the following location:
   i. Cell F4 US inflation mean 1960 -1984
   ii. Cell F5 US inflation mean 1985- 2007
   iii. Cell F6 US inflation S.D. 1960 – 1984
   iv. Cell F7 US inflation S.D. 1985- 2007

C. *Essay Questions.*

1. From **Chart 6**, what are the differences between Headline CPI inflation and Core CPI inflation?
2. From **Chart 7** and the results from Part B. 3., what implications do you get for these two periods?

III. An Exploration of the International Economic Growth

1. Go to the "world" tab in project1.xls and fill out the numbers in the blank cell for the 10 countries.

A. CIA the World Factbook.

1. Go to the course website
2. Click CIA the World Factbook
3. Go to Select a Country and click United States
4. Go to "People" → "Lie expectancy at birth" and Click
5. Find the corresponding numbers for 10 countries and fill into the "world" tab.
6. Go to "Economy" → "Oil – **Production**" and Click
7. Find the corresponding numbers for 10 countries and fill into the "world" tab.

(Note: if you cannot find the numbers for some countries, put zero for it)

B. Center for International Development

1. Go to the course website
2. Click Center for International Development
3. Click "Appendix Data Tables" under ● Barro and Lee's **Human Capita**l Updated Files.
4. Open the Excel File and search the data for 10 countries' **Average Years of School** in year 2000 for the total population aged 25 and over. Fill into the World.xls. *(Note: Russian is then named USSR)*

C. Center for International Development

1. Go to the course website
2. Click Index of Economic Freedom
3. Search 10 countries' score of Economic Freedom and fill into World tab.

D. Calculate the **correlation** *(see the note below)* between the country's GDP per capita and its health status, education status, natural resources, and economic freedom, respectively.

1. For the correlation between sample countries' GDP per capita and their life expectancy: Go to World Tab and click cell C15 and type "=Correl(C3: C12, D3: D12)" and hit "Enter".
2. Do the same thing for other correlations in cell C16, C17, C18.

E. Regression *(see the note below)*.

1. Click **help center** "?" on the right top corner. And type "slope" in the Search box. And click Slope function to see the details.
2. Go back to World Tab, in Cell E15, type "=slope(", and then highlight **C3: C12** and hold Ctrl and then highlight **F3:F12** and type")" and then hit Enter. The regression results will show up.

F. *Essay Questions.*

1. Explain the results you got from part D and E.

IV. An Exploration of the U.S. Government Budget

❖ Import the Data

1. Go to the course website and then click CBO under Data Resources
2. Click Historical Budget data and then click *The Budget and Economic Outlook: Fiscal Years 2009 to 2019*, and then click Historical Budget Data and download the Excel file into your preferred folder in your computer.
3. Click Tab "F-2" to see the budget data as a percentage of GDP. Copy the data into the "Budget" Tab of **project1.xls**.
4. Plot the US historical budget deficit and the public debt as the percentage of its GDP in the same chart. To get a smooth line chart, you might need to clear up the data file beforehand.
5. Choose column A, J, L and plot the data (save it as **Chart 8**). The chart is not beautiful because the scales of two series are very different. Therefore we need to improve it. What we can do is plot two series by two different scale (one reads from left axis, the other reads from right axis).
6. Click Series 2 (outstanding debt) directly on the line and then right click your mouse. Choose "Format data Series…" and then check the secondary axis. → Close.
7. Modify the chart to make it professional.

   G1. *Essay Question*: Briefly explain the chart you got.

   G2. *Essay Question:* Check other tabs and find out which component is the largest one in the discretionary spending and in mandatory spending. (note: Tab F-7~F-10)

Note 1 (Standard Deviation):

The standard deviation of random variables x_i is a measure of the spread of its values. It is defined as the square root of the variance. A large standard deviation indicates that the data points are far from the mean and a small standard deviation indicates that they are clustered closely around the mean $\bar{x}$.

For example, each of the three data sets (0, 0, 14, 14), (0, 6, 8, 14) and (6, 6, 8, 8) has a mean of 7. Their standard deviations are 7, 5, and 1, respectively. The third set has a much smaller standard deviation than the other two because its values are all close to 7. In a loose sense, the standard deviation tells us how far from the mean the data points tend to be.

- *Mean*: $\bar{X} \equiv \frac{1}{T}\sum_{t=1}^{T} X_t$
- *Variance*: $\mathrm{Var}(X_t) \equiv \sigma_x^2 \equiv \frac{1}{T}\sum_{t=1}^{T}\left(X_t - \bar{X}\right)^2$
- *Standard deviation*: $\sigma_x \equiv \sqrt{\sigma_x^2}$
- *Covariance*: $\mathrm{Cov}(X_t, Y_t) \equiv \sigma_{xy} \equiv \frac{1}{T}\sum_{t=1}^{T}\left(X_t - \bar{X}\right)\left(Y_t - \bar{Y}\right)$
- *Correlation*: $\mathrm{Corr}(X_t, Y_t) \equiv r_{xy} \equiv \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

To find corresponding statistical functions in Excel, you can click "Insert" → "Function" → In the select a category, choose "Statistical" and you will most statistical functions.

Note 2 (Correlation):

The correlation indicates the strength and direction of a linear relationship between two random variables. It has been suggested the following interpretations for correlations. However, all such criteria are in some ways arbitrary and should not be observed too strictly. This is because the

interpretation of a correlation coefficient depends on the context and purposes.

Correlation	Negative	Positive
Small	−0.29 to −0.10	0.10 to 0.29
Medium	−0.49 to −0.30	0.30 to 0.49
Large	−1.00 to −0.50	0.50 to 1.00

Note 3 (Regression):

$Y = a + bX$

Y is dependent variables and X is independent variables. A is intercept and b is the slope. The above regression can be interpreted as "**one unit change of X will predict *b* unit change of Y.**"

Ex: say Y is your weight (lbs) and X is the number of hours you exercise per week. Say we got *b* = -0.5. That said, if you work out one more hour per week, you will be expected to lose weight by 0.5 pound on average.

Correlation and linear regression are not the same. Consider these differences: Correlation quantifies the degree to which two variables are related. Correlation does not find a best-fit line (that is regression). You simply are computing a correlation coefficient (r) that tells you how much one variable tends to change when the other one does. With correlation you don't have to think about cause and effect. You simply quantify how well two variables relate to each other. With regression, you do have to think about cause and effect as the regression line is determined as the best way to predict Y from X. We decide the cause (X) and effect (Y) based on prior information or common knowledge. Sometimes it might not represent the true causality.

- With correlation, it doesn't matter which of the two variables you call "X" and which you call "Y". You'll get the same correlation coefficient if you swap the two. With linear regression, the decision of which variable you call "X" and which you call "Y" matters a lot, as you'll get a different best-fit line if you swap the two. The line that best predicts Y from X is not the same as the line that predicts X from Y.

Correlation is almost always used when you measure both variables. It rarely is appropriate when one variable is something you experimentally manipulate. With linear regression, the X variable is often something you experimental manipulate (time, concentration...) and the Y variable is something you measure.

Note 4 (Reference):

There are two kinds of references in Excel: relative and absolute.

Relative

For example, in Cell B2, you type "=3*A2". When you scroll down the formula for column B, you will get the corresponding formula: "=3*A3", "=3*A4", "=3*A5", "=3*A6", etc…

Absolute

If you want to fix the specific value in Row 2, you should type "3*A$2". The "$" sign fix Row 2. When you scroll down the formula for Column B, the formula is "3*A$2", "3*A$2", ..etc. However, if you scroll right the formula to the rest of Row 2, the formula will be "3*B$2", "3*C$2", "3*D$2", ….etc.

If you want to fix the specific value for Column B, you should type "=3*$B2".