Economics 304 — Fall 2017

Homework #2 — Due: 9am, Friday, September 22

Country assignments

Name	Country	Name	Country
Airi	Australia	Rebeyrol	Mexico
Dalzell	Austria	Reinitz	Netherlands
Hashmi	Canada	Richmond	New Zealand
Ilavarasan	Denmark	Samel	Norway
Lu	Finland	Szatmary	Portugal
Magleby	France	Ward	Spain
Mead	Ireland	Williams	Sweden
Nigro	Italy	Yeung	Switzerland
Piazza	Japan	Zitler	United Kingdom

Introduction

In this project, you will analyze the economic growth performance of one country over the available sample period. The sections below describe a set of analyses to be performed. You are to write up your results in a report that should include sections corresponding to the sections of the assignment.

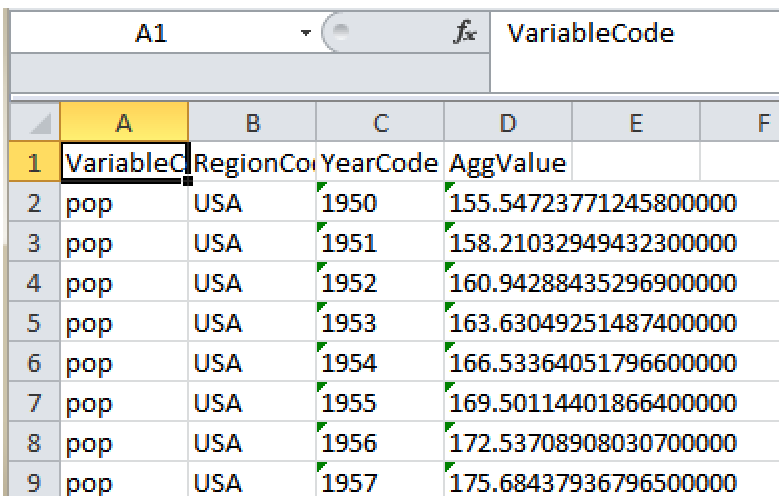
The growth analysis can be conducted completely in Excel. It could also be done in Stata, but the instructions below are aimed at using Excel. You are not expected to have any substantial familiarity with Excel, but you will learn some as you go. The instructions are very detailed and should allow you to do all of the analysis quite painlessly. (The goal here is to understand and interpret the macroeconomics of your results, not to emphasize calculation for its own sake. Please don't be offended if the instructions underestimate your software savvy.) You should submit an annotated Excel workbook/spreadsheet containing your analysis along with your report. Your report should refer to the sheets/cells in your Excel workbook that contain the results you are discussing.

The sections of this document that are shown in red are descriptions of what should go into your report.

Data

You should download the appropriate series from the Penn World Tables (http://www.rug.nl/ggdc/productivity/pwt/), which is the leading data source for internationally comparable GDP series. The current version PWT 9.0. The series are measured in PPP-equivalent real U.S. dollars, which makes them highly suitable for long-run analysis (and for inter-country comparisons, which we do not do in this project). The data are available for most countries since 1950 or 1960 and have been adjusted to the extent possible to account for changes in measurement over time.

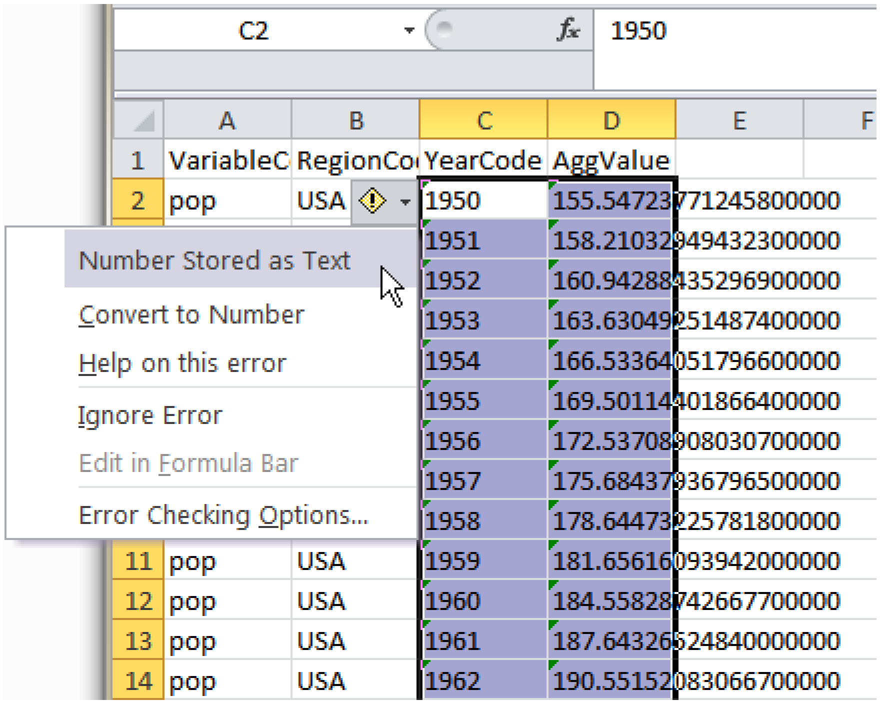
To access the data you need, click the “GA” (growth accounting) link from the “Selected data access” list. (You might need to scroll down to get the list.) Once this opens up, click “Adjust criteria” at the top. Under the Variables tab, add in population, average hours worked, and share of labor compensation if they are not already selected. Under the “Regions” tab, deselect the selected countries and choose the one you want. Then click the Results link at the top to see the table. If it’s not quite right, toggle back to “Adjust criteria” and try again. Once you get the series you want, click on “Export to .XLSX” above the data table. The data series will be on the *second tab* of the downloaded Excel spreadsheet and should look something like the figure below. (Note that my demonstrations here are from Excel for Windows. The Mac version may looks somewhat different. If you decide to use Google Sheets or some other spreadsheet program, you are on your own. The commands may be equivalent to Excel, or they may not.)



A1 | VariableCode

	A	B	C	D	E	F
1	VariableC	RegionCo	YearCode	AggValue		
2	pop	USA	1950	155.54723771245800000		
3	pop	USA	1951	158.21032949432300000		
4	pop	USA	1952	160.94288435296900000		
5	pop	USA	1953	163.63049251487400000		
6	pop	USA	1954	166.53364051796600000		
7	pop	USA	1955	169.50114401866400000		
8	pop	USA	1956	172.53708908030700000		
9	pop	USA	1957	175.68437936796500000		

When you download the data, you will probably find a small green triangle in the top left corner of the data cells. This indicates that there is an issue with the formatting of the cell, in this case the cell is formatted as text rather than as a number. You will need to change this in order to do any computations. To do so, select all of the affected data cells (most likely those in columns C and D starting in row 2). When you select one or more of these cells, an icon with an exclamation mark inside a yellow diamond should appear next to the cells. Click this icon to open a menu of choices as shown below.



Click on "Convert to Number" and the cells should become numeric, with the green triangle disappearing (and the formatting changing to right-justified rather than left-justified). You are now ready to create your dataset.

If will be easier to work with the data if you create a new spreadsheet with the variables in columns next to each other. You can copy and paste the numbers from the downloaded spreadsheet into the new spreadsheet one variable at a time. Put the variable names in the top row just to keep things straight. Your dataset should then look something like this, with no green triangles or exclamation marks:

	A	B	C	D	E	F	G	H	I	J
1	year	pop	emp	avh	hc	rgdpna	rkna	rtfpna	labsh	
2	1950	155.5635	62.835	1983.738	2.583213	2254262	7368829	0.567751	0.635634	
3	1951	158.2269	65.08094	2024.002	2.596	2435924	7653916	0.582346	0.640813	
4	1952	160.9597	65.85582	2020.183	2.60885	2535116	7973592	0.591632	0.651878	
5	1953	163.6476	66.78711	2014.5	2.621764	2654125	8318338	0.603947	0.651648	
6	1954	166.5511	65.59514	1991.019	2.634742	2639159	8650076	0.601929	0.644531	
7	1955	169.5189	67.53133	1997.761	2.647784	2827117	9021823	0.620077	0.633919	
8	1956	172.5552	69.13016	1984.025	2.662317	2887396	9390762	0.61552	0.648066	
9	1957	175.7028	69.46853	1956.493	2.67693	2948191	9743166	0.62182	0.646437	
10	1958	178.6634	68.20398	1923.115	2.691623	2926515	10053275	0.622434	0.645884	
11	1959	181.6752	69.77827	1948.32	2.706396	3128511	10434712	0.639398	0.638004	
12	1960	184.5776	70.94367	1930.096	2.721251	3208711	10803927	0.642442	0.646128	
13	1961	187.6629	71.00194	1915.404	2.755555	3290665	11151112	0.649094	0.643108	
14	1962	190.5715	72.05692	1934.786	2.790292	3491939	11569603	0.663668	0.638176	

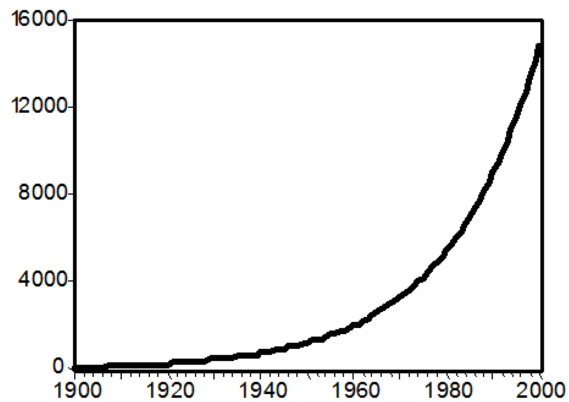
I recommend that you conduct the various steps of your analysis in separate spreadsheets (tabs) within a single Excel workbook. Name the first tab (which you have just created) Data by double-clicking on the tab (where it says Sheet1) at the bottom of the screen and typing the new name. Before starting to analyze your data, make one last check to make sure everything looks right.

Task #1: Informal assessment of real GDP growth

Begin your analysis with an informal examination of the behavior of real GDP and per-capita GDP during your sample. To do this, copy the columns for year, population (pop), and real GDP (rgdpna) to a new sheet/tab. Create a new variable for per-capita real GDP (rgdppc below) by dividing GDP by population: type =C2/B2 into cell D2, then copy cell D2 and paste it into the cells below it in column D all the way to the bottom of the table. (When you copy and paste a formula in Excel, any references to other cells in the formula are automatically translated to the new location. So when you copy the formula =C2/B2 from D2 to D3, it automatically becomes =C3/B3.) Your sheet should look something like this (but with numbers for your country, of course):

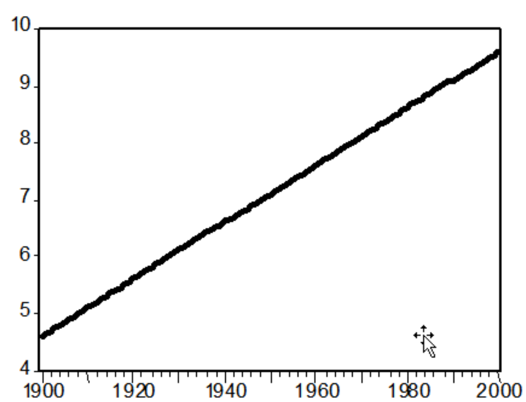
	A	B	C	D
1	year	pop	rgpdna	rgdppc
2	1950	155.5472	1979349	12725.07
3	1951	158.2103	2132526	13479.06
4	1952	160.9429	2214200	13757.68
5	1953	163.6305	2316121	14154.58
6	1954	166.5336	2301504	13820.06
7	1955	169.5011	2467224	14555.79
8	1956	172.5371	2516011	14582.44
9	1957	175.6844	2566576	14609.02

When analyzing economic growth, we use the (natural) logarithm of the variable rather than its level. A variable growing at a constant percentage rate will follow an upward-curving exponential path such as the one below:



The mathematical formula for this curve is $y(t) = y(0)e^{gt}$ with *g* being the constant rate of growth and *y*(0) the level of *y* when *t* = 0. Our eyes are not used to assessing exponential functions, so would be difficult to distinguish the constant-growth-rate diagram above from one that curved up a bit more (an increasing growth rate) or a bit less (a decreasing growth rate).

If we take the natural log of this equation, we get $\ln(y(t)) = \ln(y(0)) + \ln(e^{gt})$, or $\ln(y(t)) = a + gt,$ with *a* = ln(*y*(0)) and *g* being the growth rate. (Be sure to use the LN function in Excel, not the ~~LOG function~~.) The natural log of *y* is a linear function of time whose slope is equal to the growth rate *g*, so if we plot the log of GDP it will be a straight line if GDP is growing at a constant rate, as shown below



If the log series curves upward (downward), then the slope and growth rate are increasing (decreasing). Our eyes are good at detecting deviations from straight lines, so graphing growing time series in log form makes it easy to see how the growth rate varies over time.

Construct a time plot (line graph) of (1) the natural log of real GDP and (2) the natural log of per-capita real GDP for your country using the graph features of Excel (on the Insert menu). Based on your graphs,

- Ignoring cyclical variations, does the growth rate of each variable appear roughly constant over time or are there breakpoints at which the long-run growth rate increases or decreases?
- Are there periods of recession or boom when the economy seems to be well below or above its steady-state path?

Task #2: Estimating a growth trend

Your time plots of log real GDP and log real per-capita GDP are unlikely to be exactly straight lines. We can estimate the growth relationship over the sample period by running a *linear regression* of log GDP on time—essentially we are estimating the coefficients *a* and *g* in the equation $\log(y(t)) = a + gt$ to give the "best fit" over the sample, minimizing the squared vertical deviations from the fitted line.

Regressions are most easily done in Stata, but for this exercise we can use Excel quite easily. To estimate the slope (*g*) of the relationship between log GDP (either one) and year, click on an empty cell and use the SLOPE function. For example, if the log real GDP series is in cells E2:E63 and the years are in A2:A63, then type =100*SLOPE(E2:E63,A2:A63) to calculate the slope of the best-fit line and multiply it by 100 to put the result in percentage terms. (It will help you keep things straight if you use the cell above the one containing your calculation to type in a label telling what the number is.)

Using the full sample, what are the trend growth rates of real GDP and per-capita real GDP for your country? If your visual inspection revealed possible breakpoints where the slope of the trend line changed, perform separate trend regressions for the sub-samples before and after the break. Does there seem to be a large difference in the estimated trend growth rates before and after the break? (There are ways to test this statistically, but you would need Stata to do this so it is not required. You are welcome to do this if you know how or want to learn. Come and talk to me if you want more assistance.)

Task #3: Average vs. trend growth rates

The average (continuously compounded) growth rate over the period T_1 to T_2 is

$$\bar{g} = 100 \times \frac{\ln(y(T_2)) - \ln(y(T_1))}{T_2 - T_1}.$$

Calculate the average growth rates of real GDP and per-capita real GDP over the full available sample and compare them to the trend rate? Are they larger or smaller? Why?

Lying with statistics: Average growth rates are very sensitive to the choice of endpoints T_1 and T_2. Starting the period in the middle of a recession and ending it during a boom yields a higher average growth rate. Similarly, starting in a boom and ending in a recession gives a smaller average growth rate. If you were trying to make an optimistic case for growth in your country, what starting and ending years (using most of the years in your sample) would you choose? What if you were trying to be pessimistic? How much difference does it make?

Task #4: Convergence

One of the key implications of the Solow growth model is that countries with similar parameters (saving rate, population growth rate, production function, etc.) but different initial capital/labor ratios should eventually converge to similar levels of per-capita GDP. This implies that countries (again, with similar parameters) that start further below the leader country should grow faster than the leader.

Assess the degree of convergence for your country by calculating its initial per-capita GDP gap relative to the U.S. (early-period U.S. numbers are in the Excel demonstration tables above, or can be obtained from the PWT online) and comparing its per-capita growth rate to the U.S. rate (about 2% from 1950 to 2014). Empirical evidence suggests that each 1% gap in initial GDP typically results in a growth rate that is about 0.02% points higher, so a country that starts 20% below the U.S. should grow about 0.4% points per year more quickly than the U.S. (*i.e.*, at about 2.4% per year). How does your country measure up to the statistical average? Given its initial gap, has it grown faster or slower than would be predicted?

Task #5: Basic growth accounting and Solow residuals

One of the earliest and most fundamental modes of growth analysis is attempting to disaggregate the growth of GDP into components due to increases in the capital stock, increases in labor input, and improvements in total factor productivity—so-called "growth accounting." The "GA" dataset within the Penn World Tables was designed to facilitate growth accounting by providing time series for real GDP, employment, human capital per worker, average hours worked, and capital input.

From the data you have downloaded, you can calculate the growth rates of employment, human capital per worker, and capital input. In order to combine the labor and capital inputs we need to know how much each contributes at the margin to output. Suppose that we approximate the aggregate production function for your economy as

$$Y_t = A_t K_t^{\alpha} \left(HC_t HW_t, L_t \right)^{1-\alpha},$$

where A is an index of total factor productivity (TFP), K is capital input, HC is human capital per worker, HW is average hours worked per worker, and L is employment. Taking natural logs gives

$$\ln(Y_t) = \ln(A_t) + \alpha \ln(K_t) + (1-\alpha)\left(\ln(HC_t) + \ln(HW_t) + \ln(L_t)\right),$$

or

$$y_t = a_t + \alpha k_t + (1-\alpha)(hc_t + hw_t + l_t)$$

where lower-case letters represent the logs of the corresponding capital-letter variables.

The continuously compounded growth rate of any variable X from period $t-1$ to period t is $\ln(X_t) - \ln(X_{t-1})$, so by differencing the log equation we get an equation relating the growth rates of all of the variables:

$$(y_t - y_{t-1}) = (a_t - a_{t-1}) + \alpha(k_t - k_{t-1}) + (1-\alpha)(hc_t - hc_{t-1}) + (1-\alpha)(hw_t - hw_{t-1}) + (1-\alpha)(l_t - l_{t-1}).$$

Using your data, you can calculate year-to-year growth rates for GDP, capital, human capital per worker, and employment, but we have no direct information on TFP. The task of growth accounting is to break down total GDP growth into the four components shown in the last equation: those due to physical capital, human capital, employment growth, and growth in TFP. We calculate the latter—the Solow residual—by solving for

$$(a_t - a_{t-1}) = (y_t - y_{t-1}) - \alpha(k_t - k_{t-1}) - (1-\alpha)(hc_t - hc_{t-1}) - (1-\alpha)(hw_t - hw_{t-1}) - (1-\alpha)(l_t - l_{t-1}).$$

All we need (besides the data on the right-hand-side variables) is an estimate of the α parameter, which is the elasticity of output with respect to capital. This is approximately equal to "capital's share" of GDP, which can be measured. The PWT data provide a variable labsh, which is an estimate of labor's share, or $1 - \alpha$. The only problem is that their measure of $1 - \alpha$ is not constant but varies from year to year. So when calculating growth from year $t - 1$ to t, should we use that value of α for year t, the value for year $t - 1$, or the average of the two? The last option is probably the best, so our formula for TFP growth in period t becomes

$$(a_t - a_{t-1}) = (y_t - y_{t-1}) - \frac{\alpha_t + \alpha_{t-1}}{2}(k_t - k_{t-1}) - \left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)(hc_t - hc_{t-1})$$
$$- \left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)(hw_t - hw_{t-1}) - \left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)(l_t - l_{t-1}).$$

To perform your growth accounting exercise for your country, you should construct formulas in your Excel spreadsheet to create a table with adjacent columns measuring each of the terms in the above expression:

- GDP growth $\left(y_t - y_{t-1}\right)$,
- the contribution of capital growth to GDP growth $\left(\frac{\alpha_t + \alpha_{t-1}}{2}\left(k_t - k_{t-1}\right)\right)$,
- the contribution of human-capital growth to GDP growth $\left(\left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)\left(hc_t - hc_{t-1}\right)\right)$,
- the contribution (probably negative) of average-hours-worked growth to GDP growth $\left(\left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)\left(hw_t - hw_{t-1}\right)\right)$,
- the contribution of employment growth to GDP growth $\left(\left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)\left(l_t - l_{t-1}\right)\right)$,
- and TFP growth $\left(a_t - a_{t-1}\right)$, calculated according to the formula above. Remember that the PWT variable labsh is $1 - \alpha$, not α.

For purposes of comparison, Angus Maddison calculated growth components in the table below for several countries over time periods ending in 1987:

	GDP growth	Augmented factor input contribution	Est. contributions of other sources	Total explained growth	Unexplained growth residual
France					
1913–1950	1.15	0.48	0.10	0.58	0.57
1950–1973	5.04	2.02	1.17	3.19	1.79
1973–1987	2.16	1.24	0.30	1.54	0.61
Germany					
1913–1950	1.28	1.00	0.11	1.11	0.17
1950–1973	5.92	2.42	1.27	3.69	2.14
1973–1987	1.80	0.79	0.50	1.29	0.50
Japan					
1913–1950	2.24	1.57	0.53	2.10	0.14
1950–1973	9.27	5.44	2.53	7.97	1.20
1973–1987	3.73	2.95	0.55	3.50	0.23
Netherlands					
1913–1950	2.43	2.09	0.22	2.31	0.12
1950–1973	4.74	2.32	1.56	3.88	0.83
1973–1987	1.78	1.30	–0.06	1.24	0.54
United Kingdom					
1913–1950	1.29	0.94	0.01	0.95	0.35
1950–1973	3.03	1.76	0.52	2.28	0.73
1973–1987	1.75	0.93	0.08	1.01	0.73
United States					
1913–1950	2.79	1.53	0.41	1.94	0.83
1950–1973	3.65	2.54	0.32	2.86	0.77
1973–1987	2.51	2.55	–0.14	2.41	0.10

Maddison's "augmented factor input contribution" is roughly the sum of your capital, employment, hours, and human capital components. He includes an "estimated contributions of other sources" that you do not have, so your estimates of the Solow residual may include this component as well as his "unexplained residual."

Once you have created your table, calculate a series for the growth rates of the rtfpna variable in the PWT data (which is their measure of TFP). Does your measure match? (It should.)

Using your table, discuss which components have contributed the most on average to GDP growth in your country? Have the contributions varied over time? Assess your evidence. How does it compare to Maddison's results for your country (if it is present) and others?

Final task: Conclusion

Now that you have analyzed the growth experience of your country in detail, summarize your conclusions. What are the important characteristics of growth in this country over your sample? What, if anything, seems surprising?