

Introduction

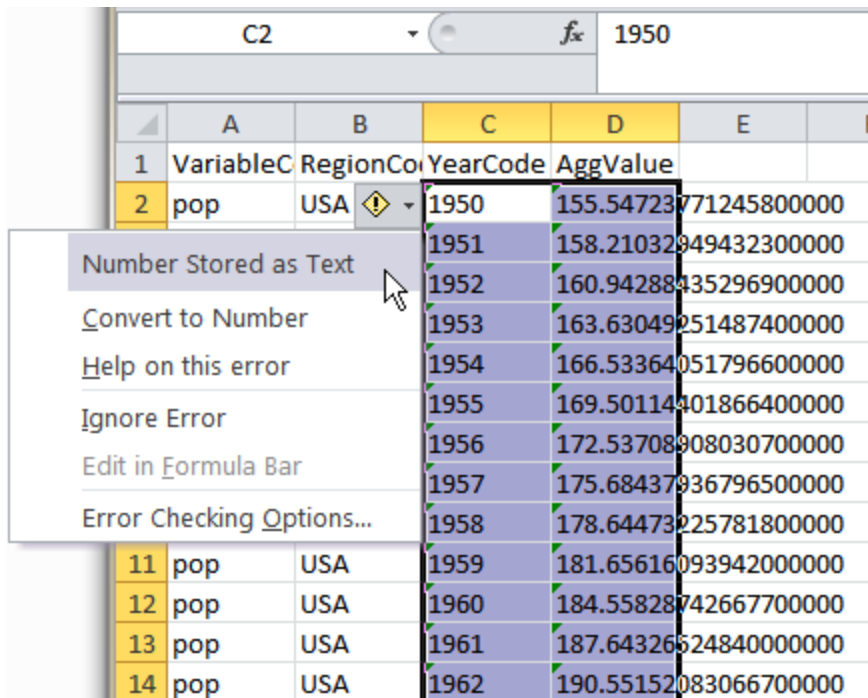
In this part of the project, you will analyze the economic growth performance of your country over the available sample period. The sections below describe a set of analyses that can be performed. Your report should roughly (or exactly) include sections corresponding to the sections of the assignment.

For this part of the project, you should download the appropriate series from the Penn World Tables (<http://www.rug.nl/research/ggdc/data/penn-world-table>), which is the leading data source for internationally comparable GDP series. The PWT series are limited to annual data, but they are measured in PPP-equivalent real U.S. dollars, which makes them highly suitable for long-run analysis. 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, select the “GA” (growth accounting) link from the list of queries. Once this opens up, click “Adjust criteria” at the top. Under the Variables tab, add in population and share of labor compensation to the ones already selected. Under the “Regions” tab, deselect the countries that are selected and choose the one(s) 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 look somewhat different.)

A1		fx VariableCode				
	A	B	C	D	E	F
1	VariableCode	RegionCode	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 that the cell is formatted as text rather than as a number. You will need to change this. 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. You are now ready to create your dataset.

It 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
1	year	pop	emp	hc	rgpdna	rkna	rtfna	labsh	
2	1950	155.5472	62.07744	2.632709	1979349	6328851	0.617148	0.639473	
3	1951	158.2103	64.5654	2.645825	2132526	6808344	0.629588	0.643262	
4	1952	160.9429	65.38007	2.659006	2214200	7050130	0.638451	0.65553	
5	1953	163.6305	66.24585	2.672253	2316121	7307389	0.651858	0.65626	
6	1954	166.5336	64.94463	2.685566	2301504	7567169	0.646179	0.649488	
7	1955	169.5011	66.76307	2.698946	2467224	7874742	0.668773	0.639841	
8	1956	172.5371	68.26872	2.713721	2516011	8177634	0.661	0.656644	
9	1957	175.6844	68.52427	2.728577	2566576	8470444	0.662165	0.655137	
10	1958	178.6447	67.29256	2.743515	2543466	8730865	0.654813	0.655887	
11	1959	181.6562	68.83398	2.758534	2725876	9042303	0.68068	0.649192	
12	1960	184.5583	69.98597	2.773636	2793428	9410647	0.678123	0.658526	
13	1961	187.6433	70.03912	2.806206	2858511	9719826	0.680692	0.65621	
14	1962	190.5515	71.20338	2.839159	3031736	10072993	0.700068	0.651404	
15	1963	193.3137	72.20612	2.872499	3164272	10465713	0.709036	0.649549	
16	1964	196.0177	73.80173	2.906231	3347373	10911239	0.723255	0.647737	
17	1965	198.4836	75.61609	2.940359	3562275	11412346	0.740149	0.642696	

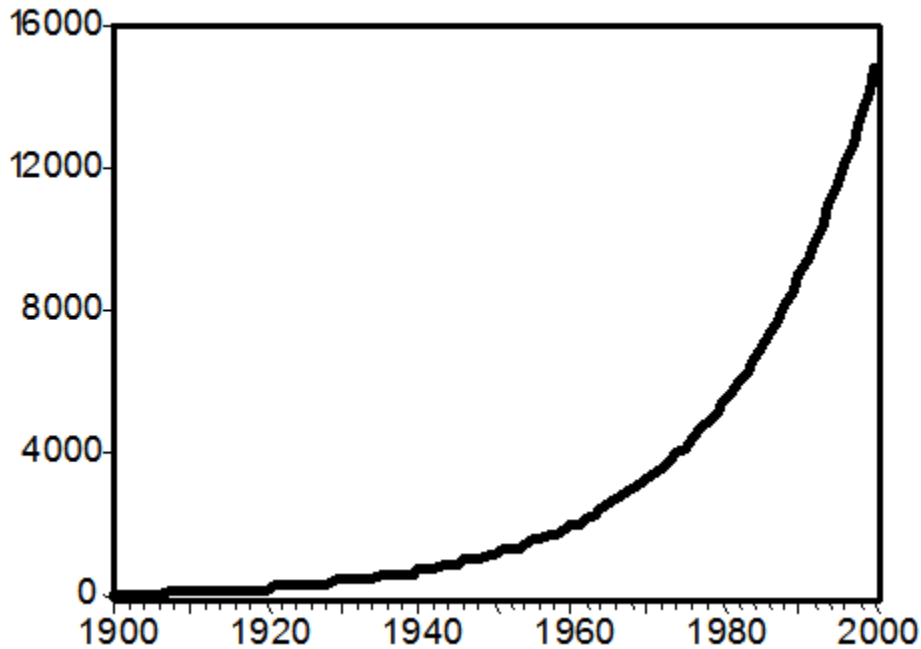
I recommend that you conduct the various steps of your analysis in separate spreadsheets (tabs) within a single Excel workbook. Name the first tab (which you have just created) Data by double-clicking on the tab (where it says Sheet1) at the bottom of the screen and typing the new name. Before starting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

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 column for year, population (pop), and real GDP (rgpdna) to a new sheet/tab. Create a new variable for per-capita real GDP 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. Your sheet should look something like this:

	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

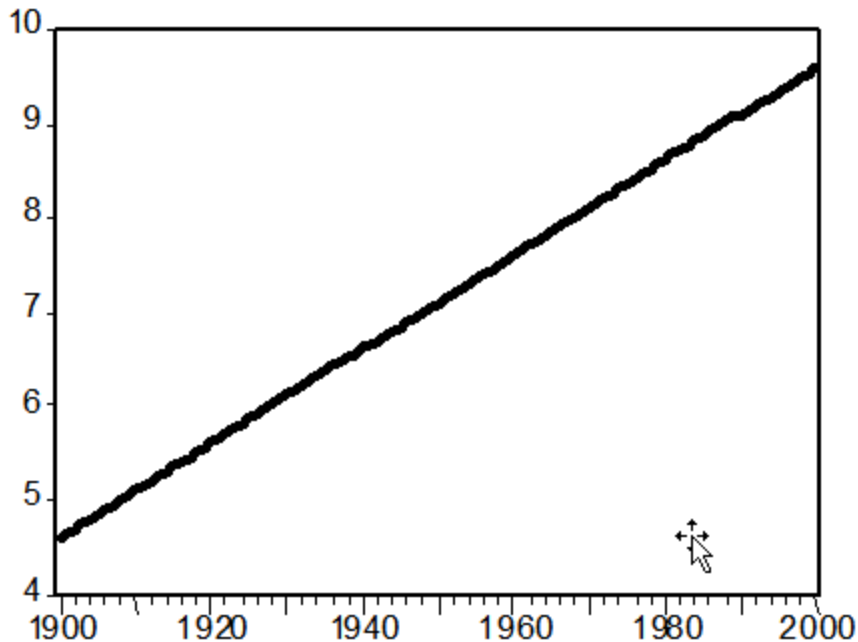
When analyzing economic growth, we use the (natural) logarithm of the variable rather than its level. A variable growing at a constant percentage rate will follow an upward-curving exponential path such as the one below:



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 log of this equation, we get $\log(y(t)) = \log(y(0)) + \log(e^{gt})$, or

$\log(y(t)) = a + gt$, with $a = \log(y(0))$ and g being the growth rate. This 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 the GDP is growing at a constant rate, as shown below



If the 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 of (1) the log of real GDP and (2) the log of per-capita real GDP for your country using the graph features of Excel (on the Insert menu). Based on your graphs,

- Does the growth rate of each variable appear roughly constant over time or are there breakpoints at which the 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?

Put the graph and your analysis into your report for this week’s project, which will become a chapter of your full analysis.

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 very easily done in Stata, but for this exercise we can use Excel equally 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 B2:B63, then type

=100*slope(E2:E63,B2:B63)

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 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 there seemed from your visual inspection that there were breakpoints where the slope of the trend line changed, perform separate trend regressions for the sub-samples before and after the break. (Include the year of the break in both sub-samples.) 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. Those of you who have taken Econ 311 or 312 are encouraged to do this. Come and talk to me to get more instructions.)

Average vs. trend growth rates

The average growth rate over the period T_1 to T_2 can be calculated as

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

Calculate the average growth rates over the full available sample and compare it to the trend rate? Are they larger or smaller? Why?

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 (within a few years of the beginning and end of the sample, respectively) would you choose? What if you were trying to be pessimistic? How much difference does it make?

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 (0.92% from 1950 to 2011). Empirical evidence suggests that each 1% gap in initial GDP 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 1.32% per year). How does your country measure up to the statistical average? Has it grown faster or slower than predicted by the empirical evidence, given its initial gap?

In the paper on the reading list by Xavier Sala-i-Martin, he finds that there are a couple of dozen variables that help explain why some countries grow faster and others more slowly than the simple convergence hypothesis would predict. Look at these variables. Based on what you know about your country, which (if any) of these variables seems like it would have an important effect in raising or lowering growth relative to the average convergence path?

Basic growth accounting and Solow residuals

One of the earliest and most fundamental modes of growth analysis is attempting to break down 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, 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

$$Y_t = A_t K_t^\alpha (H_t L_t)^{1-\alpha},$$

where A is an index of total factor productivity (TFP), K is capital input, H is human capital per worker, and L is employment. Taking logs gives

$$\log(Y_t) = \log(A_t) + \alpha \log(K_t) + (1 - \alpha)(\log(H_t) + \log(L_t)),$$

or

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

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

The growth rate of any variable X from period $t - 1$ to period t can be written as $\log(X_t) - \log(X_{t-1})$, so by differencing the log equation we get an equation for growth rates:

$$(y_t - y_{t-1}) = (a_t - a_{t-1}) + \alpha(k_t - k_{t-1}) + (1 - \alpha)(h_t - h_{t-1}) + (1 - \alpha)(l_t - l_{t-1}).$$

Using your data, you can calculate year-to-year growth rates for GDP, capital, human capital per worker, and employment, but we have no independent 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 by solving for

$$(a_t - a_{t-1}) = (y_t - y_{t-1}) - \alpha(k_t - k_{t-1}) - (1 - \alpha)(h_t - h_{t-1}) - (1 - \alpha)(l_t - l_{t-1}).$$

All we need (besides the data) 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

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

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 GDP growth $(y_t - y_{t-1})$, the contribution of capital growth to GDP growth $\left(\frac{\alpha_t + \alpha_{t-1}}{2}(k_t - k_{t-1})\right)$, the contribution of human-capital growth to GDP growth $\left(\left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)(h_t - h_{t-1})\right)$, the contribution of employment growth to GDP growth $\left(\left(1 - \frac{\alpha_t + \alpha_{t-1}}{2}\right)(l_t - l_{t-1})\right)$, and TFP growth $(a_t - a_{t-1})$, calculated according to the formula above. Using your table, which components have contributed the most on average to GDP growth in your country? Have the contributions varied over time? Assess your evidence.

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

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 different from the “standard” case that we described in class?