
Partner assignments

Amzar Abdullah	Thomas Verghese
Casey Anderson	Justin Stewart
Martis Buchholz	Brian Moore
Lauren DeRosa	Lewis May
Jess Delaney	Luis López
Jay Harman	Svetoslav Ivanov
Zachary Horváth	Sean Howard
Joseph Warren	Li Zha

The purpose of this project is to examine the phenomenon of convergence using recent growth data for many countries. As in Project #1, you will use Stata to perform some statistical analysis, but each group will work with the same data this time.

The macroeconomic topic of this project is the question of absolute and conditional convergence of per-capita incomes. There is a rich literature on this question using data from 1950 through about 1990. You should read ahead by looking at Section C of Chapter 6 of the coursebook, which reviews the standard literature on β -convergence, which discusses the Barro and Sala-i-Martin paper that you have read in addition to other literature. Your dataset and assignment picks up more or less where this earlier literature leaves off, examining the convergence question for the period 1991–2006.

The dataset for this project is called Proj3.dta. It is a Stata dataset that can be downloaded from the class Web site. You should refer to the instructions for Project #1 for basic Stata commands that we used there; the instructions in this project document assume that you are familiar with the basic commands used in Project #1. The dataset contains several growth-related variables (see Table 1) relevant to the period 1991–2006 for 174 countries. It is worth noting in passing that the data for some of these countries may be quite unreliable (Afghanistan and Iraq, for example). I have not attempted to eliminate countries with questionable data, but most of these countries have low incomes, so when you separate the sample by income you may also be screening to some extent on data quality.

As with Project #1, your results should be presented in a Word or (preferably) pdf file. The file should include all of your Stata outputs and have your names in the filename. Email the project report to parker@reed.edu by the due time.

Table 1. Variables in the Proj3.dta dataset

Variable name	Description	Units
isocode	Three-letter country code	
country	Country name	
afr	= 1 if country is in or adjacent to Africa	
westhem	= 1 if country is in the Western Hemisphere	
eur	= 1 if country is in Europe	
other	= 1 if country is not in the above groups	
ggdp	Growth rate of real GDP per capita, 1991–2006	Percentage points per year
initopen	Openness measure in 1991 ((imports + exports)/GDP)	Percentage points
initsav	Investment/GDP in 1991	Percentage points
initgdp	Real GDP per capita, 1991	Thousands of 2005 US\$
initpop	Population, 1991	Millions
initrexr	Real exchange rate in 1991	PPP with U.S. = 100, higher value indicates domestic goods are more expensive than in U.S.
aveopen	Average openness measure, 1991–2006	Percentage points
avesav	Investment/GDP averaged over 1991–2006	Percentage points
avepop	Average population, 1991–2006	Millions
averexr	Average real exchange rate, 1991–2006	PPP = 100
fingdp	Real GDP per capita, 2006	Thousands of 2005 US\$

Exercise #1: Exploring the data

- (a) Perform an eyeball test for convergence by graphing GDP growth on the vertical axis and the log of initial per-capita real GDP on the horizontal axis. Do this both for the entire sample and for each of the four regional groups: Africa, Western Hemisphere, Europe, and Other. What evidence do your graphs provide about absolute convergence of per-capita incomes during this period, both globally and regionally? Are there specific countries that are outliers?

As always, it's a good idea to hit the **browse** button on the toolbar to have a look at your data. If everything looks fine, then you can proceed.

The log of initial GDP is already created in the dataset as `linitgdp`. The GDP growth rate is `ggdp`. You can use the **Graphics** → **Twoway Graphs** selection from the menu to initiate your graph. Then Create a plot with the appropriate variables on the axes. (And you can modify your graph options as desired. If you want to get fancy, you can label each data point with the three-character variable `isocode` by playing with the **Marker properties** button.) To restrict the analysis to the African countries, select the **if/in** tab and type into the **if** box `afr`. The single-variable expression `afr` is zero

(false) or one (true) and by typing it into the box, you tell Stata to use only the observations for which this expression is one (true). You can likewise restrict the analysis to the other regional groups based on `westhem`, `eur`, and `other`. Note that there is an `if/in` tab both on the `Twoway graphs` window and on the `Plot` window. You can use either one, but if you were to type `afr` in the `Twoway if/in` window and `eur` in the `Plot if/in` window, you would exclude all observations because there are none that satisfy both criteria.

Exercise #2: Convergence regressions

- (a) Use simple regression to estimate a linear best-fit relationship between the growth rate of per-capita GDP and the log of its initial value. Do this both for the full sample and for each regional group. Use the results to test whether absolute convergence of per-capita incomes occurred for these samples.**

As discussed in Section C of Coursebook Chapter 6, the most common test for convergence is to run the following regression: $g_i = \alpha + \beta \ln(y_{0,i})$ and test the null hypothesis that $\beta = 0$. If absolute convergence occurs, then countries with lower initial per-capita income $y_{0,i}$ should have larger growth rates g_i , and the estimated value of β should be negative.

This regression can be estimated using the Stata `regress ggdp linitgdp` command. The hypothesis that $\beta = 0$ can be tested by the t statistic (and probability value) reported in the regression table. By default, Stata will use all observations that are not missing for any variable. To restrict the regression to a subset, you can, for example, append `if afr` to use only the observations for which the variable `afr` is one (true).

- (b) Repeat the absolute convergence regressions for groups of countries based on (i) level of initial income, (ii) population size, and (iii) openness to trade. Why might we expect that absolute convergence would be more likely within these sub-samples than in the full sample? Do you find absolute convergence of per-capita incomes within some or all of these groups?**

In order to break the sample down into groups by initial income, population, and openness, you'll need first to decide on the number of groups within the sample. Do you want three groups (large, medium, and small) or more than that? Having more than four or five groups is likely to make the sub-samples too small to yield reliable estimates, but having too few groups risks putting widely disparate countries into the same groups.

Once you have decided on the number of groups for each variable, you'll need to set some (arbitrary) thresholds for group boundaries. One way (but not the only way) to do so is based on percentiles of the population. For example, you can use Stata's `centile` command to calculate the 30th and

70th percentiles for the initial income, population, and openness variables: `centile initgdp initpop initopen , centile (30 70)`. (Do not include the trailing period in this command, but be sure to include the comma: options in Stata always follow a comma.) Countries with a value below the 30th percentile are in the lowest 30 percent of the sample; those above the 70th percentile value are in the highest 30 percent. Suppose that the 30th percentile for `initgdp` is 2. (The actual value you should use will be taken from your centile output and it won't be exactly 2.) You could construct a dummy variable for low initial income by generate `lowinc = initgdp < 2`. This variable will have the value one for countries with `initgdp` less than 2 and zero for others. It can be used in an "if clause" of a regress command to restrict the regression to low-income countries just as you used the regional dummy variables in the previous part. (You don't really need to construct the dummies unless you want to; you could just type `reg ggdpc linitgdp if initgdp < 2` to put the condition in the regress command directly and restrict the sample to those countries with 1991 per-capita GDP below \$2,000.)

- (c) Test for conditional convergence by including additional variables in the regression that might affect the position of a country's steady-state growth path. What effects would you expect each of these variables to have on the level of a country's steady-state growth path? What other variables (not in your dataset) would you like to have? What do your results say about conditional convergence in the sample period?**

Breaking up the sample into subsamples is one way of taking account of level differences in countries' steady-state growth paths. Another way is to use "conditional-convergence" regressions by including determinants of the growth path as additional explanatory variables alongside the log of initial per-capita income.

As discussed in Coursebook Section 6.C, empirical evidence in the earlier post-World War II sample is more favorable to conditional than to absolute convergence. Your dataset includes several variables that have been argued to affect the level of the steady-state growth path: the saving (investment) rate, the degree of openness of the economy, and the real exchange rate (terms of trade).

Each of these variables is available in two forms: its average over the period and its value at the beginning of the period. While the average can be argued to be more representative of the entire period, it might lead to econometric problems if there is "reverse causality" from growth to these variables. The initial values are arguably exogenous and thus more acceptable as right-hand-side regressors, but they are less representative of the entire period. Thus, there is no obvious choice as to which set of variables is better to use.

To test for conditional convergence, rerun the convergence regressions for the full sample including these variables as additional regressors, first using the average values then using (in a separate regression) the initial values. Does the coefficient on initial per-capita income now show evidence of convergence? How about for the regional groups?

- (d) Summarize your results on tests of convergence for the 1991-2006 period. To what extent do they agree or disagree with other studies in the literature?**