Economics 314 Spring 2010
Project #3 Assignment Due: 9am, Wednesday, February 17

Partner assignments

Axcelle Bell	Nisma Elias
Niall Murphy	Gabriel Korzeniewicz
Nicholas Wilson	Antonio Friedman
Zakir Hakim	Mrijan Rimal
Mark Hintz	Suraj Pant
Samantha Bruce	Ximeng Zhang

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. Your dataset and assignment picks up where this earlier literature leaves off, examining the convergence question for 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	US$ = 100, higher value indicates domestic goods are more expensive
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	US$ = 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-capital 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: Absolute and conditional 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: $g_i = \alpha + \beta \ln\left(y_{0,i}\right)$ 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. Be 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 (initially) relatively rich countries, middle-income countries, poorer countries, relatively big (population-wise) countries, middle-sized countries, and smaller countries. Did per-capita incomes in these groups converge absolutely?

Setting thresholds for high, middle, and low with respect to income or population is necessarily arbitrary. One way to do so is based on percentiles of the population. Use Stata's centile command to calculate the 30th and 70th percentiles for the initial income and population variables: centile initgdp initpop , centile (30 70). (Do not include the trailing period in this command, but be sure to include the comma.) Countries 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 variables for low-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 `if initgdp < 2` in the regress command directly.)

(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.

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, 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.

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, size groups, and initial-income groups?

(d) Summarize your results on tests of convergence for the 1991–2006 period. To what extent to they agree or disagree with those from earlier periods?