Homework 2: ...But We Make It Up in Volume

36-402, Spring 2015

Due at 11:59 pm on Monday, 26 January 2015

"Gross domestic product" is a standard measure of the size of an economy; it's the total value of all goods and services bought and solid in a country over the course of a year. It's not a perfect measure of prosperity[1], but it is a very common one, and many important questions in economics turn on what leads GDP to grow faster or slower.

One common idea is that poorer economies, those with lower initial GDPs, should grower faster than richer ones. The reasoning behind this "catching up" is that poor economies can copy technologies and procedures from richer ones, but already-developed countries can only grow as technology advances. A second, separate idea is that countries can boost their growth rate by under-valuing their currency, making the goods and services they export cheaper.

This week's data set contains the following variables:

- Country, in a three-letter code (see `http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3`).
- Year (in five-year increments).
- Per-capita GDP, in dollars per person per year ("real" or inflation-adjusted).
- Average percentage growth rate in GDP over the next five years.
- An index of currency under-valuation[2]. The index is `0` if the currency is neither over- nor under- valued, positive if under-valued, negative if it is over-valued.

Note that not all countries have data for all years. However, there are no missing values in the data table.

[1] A standard example: if vandals break all the windows on a street, a town, GDP goes *up* by the cost of the repairs.

[2] The idea is to compare the actual exchange rate with the US dollar to what's implied by the prices of internationally traded goods in that country — the exchange rate which would ensure "purchasing power parity". The details are in the paper this assignment is based on, which will be revealed in the solutions.

1. (10) Linearly regress the growth rate on the under-valuation index and the log of GDP. Report the coefficients and their standard errors (to reasonable precision). Do the coefficients support the idea of "catching up"? Do they support the idea that under-valuing a currency boosts economic growth?

2. (20) Repeat the linear regression but add as covariates the country, and the year. Use `factor(year)`, not `year`, in the regression formula.

   (a) (5) Report the coefficients for log GDP and undervaluation, and their standard errors, to reasonable precision.

   (b) (5) Explain why it is more appropriate to use `factor(year)` in the formula than just `year`.

   (c) (5) Plot the coefficients on year versus time.

   (d) (5) Does this expanded model support the idea of catching up? Of under-valuation boosting growth?

3. (10) Does adding in year and country as covariates improve the predictive ability of a linear model which includes log GDP and under-valuation?

   (a) (1) What are the R^2 and the adjusted R^2 of the two models?

   (b) (5) Use leave-one-out cross-validation to find the mean squared errors of the two models. Which one actually predicts better, and by how much? *Hint:* Use the code from lecture 3.

   (c) (4) Explain why using 5-fold cross-validation would be hard here. (You don't need to figure out how to do it.)

4. (20) *Kernel smoothing* Use kernel regression, as implemented in the `np` package, to non-parametrically regress growth on log GDP, under-valuation, country, and year (treating year as a categorical variable). *Hint:* read chapter four carefully. In particular, try setting `tol` to about 10^{-3} and `ftol` to about 10^{-4} in the `npreg` command, and allow several minutes for it to run. (If you are using R Markdown, trying caching this part of your code.)

   (a) (5) Give the coefficients of the kernel regression, or explain why you can't.

   (b) (5) Plot the predicted values of the kernel regression, for each country and year, against the predicted values of the linear model.

   (c) (5) Plot the residuals of the kernel regression against its predicted values. Should these points be scattered around a flat line, if the model is right? Are they?

   (d) (5) The `npreg` function reports a cross-validated estimate of the mean squared error for the model it fits. What is that? Does the kernel regression predict better or worse than the linear model with the same variables?

5. (20) *Time courses and interactions* In this question, use the kernel regression you fit in the previous problem.

   (a) (6) Plot the predicted growth rate, as a function of the year, in five year increments from 1955 to 2000, if the initial GDP (not log GDP!) is $10,000 in each period, the under-valuation index is 0 (i.e., no under- or over- valuation), and the country is Turkey.

   (b) (3) Re-do the plot but change the under-valuation index to +0.5.

   (c) (3) Re-do the plot but hold the initial GDP at $1,000 and the under-valuation index at 0.

   (d) (3) Re-do the plot with the initial GDP at $1,000 and the under-valuation index at +0.5.

   (e) (5) Is there evidence of an interaction between initial GDP and under-valuation? Explain.

6. (20) *Average predictive comparisons* Section 4.5 of the notes explains how to calculate the "average predictive comparison" — the typical rate of change in the response when a given variable is perturbed, even when the model is nonlinear and has interactions. See, in particular, Equation 4.31.

   Hint: at no point in this problem should you re-fit either model.

   (a) (5) Calculate the average predictive comparison for log GDP in the kernel regression.

   (b) (5) Calculate the average predictive comparison for under-valuation in the kernel regression.

   (c) (5) Explain how to calculate the corresponding average predictive comparisons from the linear model's coefficients. What are the average predictive comparisons for initial log GDP and for under-valuation in the linear model?

   (d) (5) Do the kernel and the linear regression agree, qualitatively, about the average effect of increasing initial GDP on growth? Do they agree, qualitatively, about the effect of undervaluation on growth?