PLSC 503: "Multivariate Analysis for Political Research"

Exercise One
January 26, 2017

Introduction

This homework is an opportunity to demonstrate your mastery of bivariate OLS regression. The exercise contains a simulation component and a substantive part, the topic of which is the relationship between political and economic factors and HIV/AIDS.

Exercise: Simulation

For the case of a bivariate regression $Y_i = \beta_0 + \beta_1 X_i + u_i$ where $u_i \sim$ i.i.d. $N(0, \sigma^2)$, show via simulation that:

1. the OLS estimator $\hat{\beta}_1$ is an unbiased estimate of β_1,

2. $\text{Var}(\hat{\beta}_1)$ is increasing in σ^2,

3. $\text{Var}(\hat{\beta}_1)$ is decreasing in N,

4. $R^2_{adj.} \to R^2$ as $N \to \infty$.

Hint: One fake regression doesn't "show" anything.

Exercise: Data Analysis

The data are from 2007, and contain information on 111 nations in the world system. In addition to identifiers, they contain four variables, and reside in .csv format on the course GITHUB repo. The variables there are:

- **HIVPrevalence** is the percentage of the people aged 15 or over in the country currently infected with the HIV virus (source: WHO / UNAIDS).

- **HIVDeathRate** is the number of people *per 1000 infected* who died from complications due to HIV/AIDS in that country (source: WHO / UNAIDS).

- **GDPPerCap** is the country's gross domestic product per capita, in constant 2005 U.S. dollars.

- **CivilWarLag** is a dichotomous variable, coded 1 if the country experienced a civil conflict in 2006 (the year prior to the HIV/AIDS measures) and 0 if it did not.

Your assignment is to use OLS regression to examine the **(bivariate)** relationships among HIV prevalence and death rates, wealth, and civil conflict, and to prepare a short report on that relationship as if you were writing up the results for a paper or journal submission. You should think of the two HIV/AIDS variables as your two "dependent" variables, and the other two as "independent" variables / covariates. Your report should contain *at least* the following:

1. Very brief summaries of the dependent and independent variables, including a graphical presentation of the relationships;

2. Results of your OLS estimate(s), including coefficients, standard errors, `SEE`s, R^2s, and other relevant statistics;

3. A discussion, in words, of those results/findings, including matters relating to marginal effects, statistical inference, and model fit.

This homework is due *electronically* by 5:00 p.m. EST on **Thursday, February 2, 2017**, and is worth 50 points.