

Quant II - Problem Set IV

Diagnostics

Kenneth Benoit

Assigned: Wednesday, January 28, 2009

Due: Wednesday, February 4, 2009

1. Tavits (2005) examines electoral volatility in post-communist countries.
 - (a) Load the STATA-dataset and regress the dependent variable (volatility) on ENP, part_age, dismag, polarization, GDP, inflation, heterogeneity, urb_rur, yearsdem, and yearsdem2. Showing the command and the table as printed by R is sufficient, there is no need for discussing the results at this stage.
 - (b) Create a residual-versus-fitted plot and discuss the findings in a few sentences. Make sure you answer the following questions: What does the red line tell you about the model's performance in predicting electoral volatility? What does the plot (as such, no need to do additional things) tell you about the numbered observations in *substantive* terms? Imagine you would have to explain this to a friend who knows very little about statistics.
 - (c) Perform an F-test to check whether the variances of the residuals are the same in two samples split at the median (referring to the regression sample) of the fitted values. Briefly discuss the result. What is your conclusion?
 - (d) Plot Cook's d against the observation numbers. Which are the country-elections with the three highest values? What does it mean if these cases have high Cook's d values?
2. A closer look at the link between inflation and volatility.
 - (a) Referring to the sample used in the regression, give descriptive statistics for the variable inflation and create a kernel density plot. What do you see?
 - (b) What is the substantive assumption of including inflation as it is in the regression?
 - (c) Having this assumption in mind, briefly interpret the findings for inflation in the above regression.

- (d) Suggest a meaningful transformation of the variable inflation in order to avoid the problematic assumption. (Note: You may need to add a small constant to the variable before transforming it, so that all values to be transformed are positive and you do not lose additional cases.)
- (e) Re-run the model from above while replacing inflation with its transformed variant. Compare the “new” finding with regard to inflation with the one from above.