<u>Problem Set 7</u>: More Interactions and Omitted Variables

I. INTERACTIONS: Consider a pairwise linear-interactive linear model of the proposition that Government Duration is a function of the number of parties in government, parliamentary support for government, and party discipline. *Pairwise linear-interactive* here means a model that entertains hypotheses that the effects of each variable on government duration depends on each of the others, but not on any particular combination of the others (i.e., just first-order or pairwise and no higher-order interactions).

$DGOVPW=\beta_0+\beta_1 PSUPGPW+\beta_2 NPGOVPW+\beta_3 PD+\beta_4 PD*PS+\beta_5 PD*NP+\beta_6 NP*PS+\varepsilon$

A. Estimate this model using our data set.
B. What is the estimated effect of parliamentary support on government duration?
C. What is the estimated effect of the number of parties in government on government duration?
D. What is the estimated effect of party discipline on government duration?

Using any graphics program you like (a spreadsheet is convenient, or any of the various Stata code mentioned in class, or other code, including your own)...

F. Create a graph of the effect of parliamentary support on government duration, as a function of the number of parties in government, when party discipline is high. Create another graph of the same thing, only when party discipline is low. On both of these graphs, graph a 90% confidence interval around the effect line as well. (You may wish to put both of these plots on one graph also, or instead.)
G. Graph the effect of the number of parties in government on government duration, as a function of parliamentary support for the government, when party discipline is high. Create another graph of the same thing, only when party discipline is low. On both of these graphs, have a 90% confidence interval around the effect line graphed as well. (You may wish to put both plots on one graph.)
H. Create a table of the effects of party discipline at various levels of parliamentary support and number of parties in government. In particular, parliamentary support (PSUPG) runs approximately 40 to 80 (%) in our sample (i.e., its sample range is 40-80) and the number of parties runs from 1 to about 4. The mean of parliamentary support is about 57 and the mean number of parties is about 2. So, create a 3x3 table of the effects of party discipline (measured by the movement from a low PD country to a high one) on government duration at the various combinations of these values for the other variables. Each cell entry in the table should include the effect of PD at that combination of PS and NP, the standard error of that effect, and a t-test that the effect at that point is zero.

II. OMITTED VARIABLES: Return to this model from Problem Set 5:

(1) $SSPENDG = \beta_c + \beta_a AGE + \beta_u UE + \beta_n NPGOVPW + \beta_v VPART + \varepsilon$.

Suppose (1) were the true model, but you instead estimate (2). (Alternatively, suppose you wish to evaluate some critic(s)'s claim that you omitted VPART in estimating (2) and should have estimated (1)):

(2) $SSPENDG = \beta_c^* + \beta_a^* AGE + \beta_u^* UE + \beta_n^* NPGOVPW + \varepsilon$

A. What is β_n^* in terms of (a) coefficient(s) of model (1) and of some auxiliary regression? Estimate the regression necessary to estimate b_n^* (*i.e.* the estimate of β_n^*), whatever coefficient(s) you needed from model (1) to answer the first part of this question, and whatever other thing(s) you may have needed. Show that b_n^* is indeed equal to the estimated analog of what you said β_n^* was equal to.
B. Under what conditions would b_n^* be an unbiased estimate of β_n, the effect of NP on SS spending controlling voter participation, age, and unemployment? *I.e.*, when would $E(b_n^*) = \beta_n$?
C. Describe the conditions under which the s.e.(b_n^*) be less than, greater than, or equal to the s.e.(b_n)? Under what conditions would b_n^* be efficient (have minimum variance)?

Suppose, now, (1) is still the truth but you estimate: $SSPENDG = b_c + b_a AGE + b_u UE + b_n NP + b_v VP + b_p PD + \varepsilon$.

D. What is the expected value of b_p, $E(b_p)$? What about the other coefficients?
E. Are your coefficient estimates efficient? Why or why not?