PLSC 503: Problem Set 5

Thad Dunning

Department of Political Science

Yale University

Spring 2010

Due February 16, 2010

1 Theoretical/Conceptual Exercises

1. **Interaction terms**. Scholars sometimes counsel the use of interaction models for testing conditional hypotheses. This exercise gives an example. Assume the following OLS model:

$$Y_i = a + bX_i + cZ_i + dX_iZ_i + \epsilon_i \tag{1}$$

for all $i = 1, ..., n$. Here, X_iZ_i is the product of X_i and Z_i; this is usually called an "interaction term." The usual OLS assumptions apply, e.g., $\epsilon_i \perp\!\!\!\perp (X_i, Z_i, X_iZ_i)$ and the ϵ_i are i.i.d. with $E(\epsilon) = 0$ and $\text{var}(\epsilon_i) = \sigma^2$ for all i.

   (a) What is the marginal effect of intervening to change X_i with Z_i held fixed? That is, if

$$E(Y_i|X_i, Z_i) = a + bX_i + cZ_i + dX_iZ_i, \tag{2}$$

   what is

$$\frac{\partial E(Y_i|X_i, Z_i)}{\partial X_i}? \tag{3}$$

   And what is the marginal effect of intervening to change Z_i with X_i held fixed?

   (Note: here, the "marginal effect" means, the change in the expected value of Y_i due to a one-unit change in X_i or Z_i).

   (b) Now, suppose you run a regression of Y_i on a constant, X_i, Z_i, and X_iZ_i, obtaining

$$\hat{Y}_i = \hat{a} + \hat{b}X_i + \hat{c}Z_i + \hat{d}X_iZ_i \tag{4}$$

   for all i. Here, $\hat{Y}_i$ is the fitted ("predicted") value of Y_i, $\hat{a}$ is the fitted intercept, and $\hat{b}$, $\hat{c}$, and $\hat{d}$ are the fitted slope coefficients. Under the assumptions of the model in (1), $\hat{a}$ estimates a, $\hat{b}$ estimates b, and so on.

What is the estimated marginal effect of intervening to change X_i with Z_i held fixed? How about the estimated marginal effect of intervening to change Z_i with X_i held fixed?

(c) Let $\frac{\partial \hat{Y}_i}{\partial X_i}$ be the estimated marginal effect of intervening to change X_i with Z_i held fixed. Express the variance of $\frac{\partial \hat{Y}}{\partial X_i}$ in terms of variances and covariances of the variables and coefficient estimators in equation (4). (For convenience, treat X_i and Z_i as fixed, rather than as random variables).

(d) Suppose that after fitting equation (4), you find that

$$\hat{Y}_i = 1.2 + 2.3X_i + 0.5Z_i - 2.1X_iZ_i. \tag{5}$$

Moreover, the estimated variance-covariance matrix of $\hat{\beta} = (\hat{a}\ \hat{b}\ \hat{c}\ \hat{d})'$ is given by

$$\widehat{\text{cov}}(\hat{\beta}|X, Z, XZ) = \begin{pmatrix} 0.5 & 0.3 & 0.4 & 0.7 \\ 0.3 & 0.9 & 0.5 & -0.3 \\ 0.4 & 0.5 & 0.4 & 0.3 \\ 0.7 & -0.3 & 0.3 & 0.7 \end{pmatrix} \tag{6}$$

i. Conduct a test of the null hypothesis that $b = 0$.

ii. Conduct a test of the null hypothesis that $d = 0$.

iii. Conduct a test of the null hypothesis that the marginal effect of intervening to change X_i, with Z_i held fixed at $Z_i = 1$, is zero.

iv. Conduct a test of the null hypothesis that the marginal effect of intervening to change X_i, with Z_i held fixed at $Z_i = 10$, is zero.

v. Comment on your results. Do they suggest that—given the model—the effect of X_i is conditional on Z_i?

2. **Effect of treatment on Compliers**. Work exercise 21 in the discussion questions in Freed-

man (2009), Chapter 4 and then look at the data in the “answers to exercises” question on p. 252.

First, answer these questions. Which subjects are the “Always-Treats” in the control group? Which subjects are the “Never-Treats” in the invitation (assigned-to-treatment) group?

Now, estimate the effect of treatment on Compliers. (Show your work). Compare your results to the intention-to-treat analysis. If intention-to-treat analysis shows zero effect of treatment assignment, can the estimated effect of treatment on Compliers be non-zero? Why or why not?

2 Computer exercises

1. Open the yule.dta file (if you are working in Stata, or yule.csv if you are working with another package). Convert the four variables in the Yule dataset (Paup, Out, Old, and Pop) to percentage changes by subtracting 100 from each observation of each variable. (I will refer to these converted variables as DeltaPaup, DeltaOut, DeltaOld, and DeltaPop).

   Then generate a new variable which is the product of DeltaOut and DeltaOld; I’ll call this variable OutOld.

2. Now, use the canned command in Stata (or other package) to regress DeltaPaup on DeltaOut, DeltaOld, OutOld, and DeltaPop. (For purposes of step 3 below, it may be useful to put the independent variables in that order).

   Report your output. What kinds of hypotheses do you think that fitting this model might be useful for testing? What do the results seem to show?

3. Next, create a twoway plot in Stata that shows the marginal effect of DeltaOut on DeltaPaup, at different values of DeltaOld. Your plot should also include a 95% confidence interval

around the line plotting the marginal effect. Turn in your plot.

How to do this:

Interaction models are reviewed in: Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14: 63-82.

On their useful website relating to this article, the authors have posted Stata code that creates graphs showing marginal effects, as well as confidence intervals for the marginal effects. See http://homepages.nyu.edu/ mrg217/interaction.html.

This code is posted under the Resources tab of the classes v2 website, with the title "singlemodifyingvariable.do." I made a few small additions/notes to the code. You may want to play with the format, axis labels, etc... of the graph.

4. To what extent are the effects of changes in the outrelief ratio (that is, DeltaOut) on changes in pauperism (that is, DeltaPaup) conditional on the change in the proportion of the population that is old (that is, DeltaOld)? Draw some conclusions.