Problem Set 3

EE426 Econometrics 2

Due April 3, 2014

Please report the regression results in each problem and print STATA .do file attached at the end of your answer.

1. Consider the following three-equation structural model

$y_1 = \gamma_{12}y_2 + \gamma_{13}y_3 + \delta_{11}z_1 + \delta_{13}z_3 + u_1$ (1)

$y_2 = \gamma_{21}y_1 + \delta_{21}z_1 + u_2$ (2)

$y_3 = \delta_{31}z_1 + \delta_{32}z_2 + \delta_{33}z_3 + \delta_{34}z_4 + u_3$ (3)

Please show whether equation (1) and equation (2) can be identified using order and rank conditions. If they are identified (or either of them), please specify what type of identification (just identified or overidentified) and what variables can be used for instrumenting.

2. Use the data in OPENNESS.dta for this exercise.

Romer (1993) proposes theoretical models that more open countries should have lower inflation rates. He has a two-equation system:

$inf = \beta_{10} + \alpha_1 open + \beta_{11} log(pcinc) + u_1$ (4)

$open = \beta_{20} + \alpha_2 inf + \beta_{21} log(pcinc) + \beta_{22} log(land) + u_2$ (5)

His hypothesis is that $\alpha_1 < 0$. The second equation reflects the fact that the degree of openess might depend on the average inflation rate, and other factors. The idea is that, ceteris paribus, a smaller country is likely to be more open, $\beta_{22} < 0$.

(2.1) Check whether which equation is identified. Then, write and estimate the reduced form equation, and test whether log(*land*) can be used as an IV for the model. If so, estimate equation (4) using log(*land*) as an IV for *open*.

(2.2) Because log(pcinc) is insignificant in (2.1)'s estimation, drop it from the analysis. Estimation equation (4) by OLS and IV without log(pcinc). Do any important conclusions change?

(2.3) Return to equation (4). Add the dummy variable *oil* to the equation and treat it as exogenous. Estimate the equation by IV. Does being an oil producer have a ceteris paribus effect on inflation?

3. Use SMOKE.dta for this exercise

A model to estimate the effects of smoking on annual income (through lost work days due to illness, or productivity effects) is

$log(income) = \beta_0 + \beta_1 cigs + \beta_2 educ + \beta_3 age + \beta_4 age^2 + u_1$, (6)

where *cigs* is a number of cigarattes smoked per day, on average.

Cigarette consumption might be jointly determined with income, a demand for cigarettes equation is

$$cigs = \gamma_0 + \gamma_1 log(income) + \gamma_2 educ + \gamma_3 age + \gamma_4 age^2 + \gamma_5 log(cigpric) + \gamma_6 restaurn + u_2, \quad (7)$$

where *cigpric* is the price of a pack of cigarettes (in cents), and *restaurn* is a binary variable equal to 1 if the person lives in a state with restaurant smoking restrictions. Assuming these are exogenous to the individual.

(3.1) How can we ceteris paribus interpret each equation? What signs would you expect for γ_5 and γ_6 ?

(3.2) Under what assumption is the income equation (6) identified?

(3.3) Estimate the reduced form for *cigs* (regressing *cigs* on all exogenous variables). Can we use log(*cigpric*) and *restaurn* as instruments?

(3.4) Estimate the income equation by 2SLS.

(3.5) Estimate the income equation by OLS. Discuss the meaning of β_1 and how the estimate of β_1 by OLS compares with β_1 by 2SLS.