Exam #2

Economics 435: Quantitative Methods

Spring 2005

Question 1 is worth 20 points, question 2 is worth 30, and question 3 is worth 50.

1 Some short-answer questions

These applied questions are long, but the (correct) answers are very short.

a) Suppose that I'm estimating a regression that aims to measure the returns to education. My dependent variable is wage at age 40, and my explanatory variables are years of education, sex, and number of children. If I am concerned that education and number of children are endogenous, how many instrumental variables do I need to consistently estimate the structural model? Is it a problem if I have more?

b) It is common to use data on siblings to "control" for unobserved family-specific influences. The setup is very similar to that for panel data with fixed effects. For example, suppose we have data on education levels and wages at age 40 for 1,000 pairs of brothers. For each pair we calculate the difference in wages between the younger and older brother, and the difference in years of education. We estimate an OLS regression of the difference in wages on the difference in education and interpret the result as measuring the effect of education on wages. Other than measurement error (since we've already talked about that), identify one potential problem with this approach. Is the problem you have identified likely to result in over or under estimation of the returns to schooling under this approach?

2 An alternative to the IV estimator

Imagine that you are helping a friend - we'll call him Clarence - with his 435 paper. Clarence has a data set $\{(y_i, x_i, z_i)\}_{i=1}^n$ generated by a random sample from the joint distribution of (y, x, z, u). The relationship among the variables is given by the structural model:

$$y = \beta_0 + \beta_1 x + u$$

where β_1 is the effect of x on y, and u is the effect of all other variables on y. We are not able to assume that $E(u|x) = 0$, but we are able to assume that z is a valid instrument:

$$\begin{aligned} cov(u, z) &= 0 \\ cov(x, z) &\neq 0 \end{aligned}$$

Finally, we assume that there is variation in both x and z in Clarence's sample.

Clarence didn't pay attention in class and has never heard of the IV or 2SLS estimator. Instead of using an IV approach, he settles on a "run every possible OLS regression" approach to his paper. He estimates three regressions by OLS: one regression of y on x, one of y on z, and one of x on z. Let the coefficients from those regressions be defined as:

$$\begin{aligned}\hat{\gamma}_{yx} &= \frac{c\hat{o}v(y,x)}{v\hat{a}r(x)}\\ \hat{\gamma}_{yz} &= \frac{c\hat{o}v(y,z)}{v\hat{a}r(z)}\\ \hat{\gamma}_{xz} &= \frac{c\hat{o}v(x,z)}{v\hat{a}r(z)}\end{aligned} \tag{1}$$

He has brought you the results from these regressions. Unfortunately, he has accidentally deleted the source data and cannot estimate any new regressions.

a) Is there a consistent estimator of β_1 that can be constructed using $(\hat{\gamma}_{yx}, \hat{\gamma}_{yz}, \hat{\gamma}_{xz})$? If so, find it and prove that it is consistent.

b) How does your candidate estimator relate to the standard IV estimator?

3 Simultaneous equations

Suppose that the market price and quantity for a particular good are determined by the intersection of a supply curve:

$$q = \lambda_0 + \lambda_1 p + \lambda_2 x + u$$

with a demand curve:

$$q = \gamma_0 + \gamma_1 p + \gamma_2 z + v$$

where the λ and γ terms are structural parameters, q is log quantity, p is log price, x is an exogenous variable that shifts supply and z is an exogenous variable that shifts demand. We assume that $E(u|x,z) = E(v|x,z) = 0$.

Suppose you have a data set (p_i, q_i, x_i, z_i) describing the market in various locations at various points in time, and that this data set can be considered a random sample.

a) In general, can we learn anything about the structural parameters from an OLS regression of q on p? Briefly and intuitively explain why this is.

b) Find p as a function of (x, z, u, v) and the structural parameters (Note: you have two equations above and two unknowns (p and q); use standard algebra to solve).

c) Find q as a function of (x, z, u, v) and the structural parameters.

d) Suppose that you have used your data to estimate an OLS regression of p on (x, z). Let $\hat{\beta}_{px}$ be the coefficient on x and $\hat{\beta}_{pz}$ be the coefficient on z from that regression. Find:

$$\begin{aligned}\beta_{px} &= \text{plim } \hat{\beta}_{px}\\ \beta_{pz} &= \text{plim } \hat{\beta}_{pz}\end{aligned} \tag{2}$$

in terms of the structural parameters.

e) Suppose that you have used your data to estimate an OLS regression of q on (x, z). Let $\hat{\beta}_{qx}$ be the coefficient on x and $\hat{\beta}_{qz}$ be the coefficient on z from that regression. Find:

$$\begin{aligned} \beta_{qx} &= \text{plim } \hat{\beta}_{qx} \\ \beta_{qz} &= \text{plim } \hat{\beta}_{qz} \end{aligned} \tag{3}$$

in terms of the structural parameters.

f) Use your results above to construct consistent estimates of the price elasticity of supply λ_1 and the price elasticity of demand γ_1 based on the reduced form coefficients $(\hat{\beta}_{px}, \hat{\beta}_{pz}, \hat{\beta}_{qx}, \hat{\beta}_{qz})$