Econ. 240B, Second Half — Daniel McFadden, 1999

**EXERCISE 3. GROWTH AND EXIT OF FIRMS (continued)**
(To be handed in on Nov. 16)

This exercise uses the data on exit of firms from Exercise 1. As before, 8031 observations can be found in the file ex1-1.dat in format ‘f12.0,3f10.0,3g11.0,f9.0,11g11.0', with the variables:

| | |
|---|---|
| id | firm id |
| year | 4 digit year, between 1986 and 1995 |
| sic | 4 digit sic code |
| ind | 2 digit industry code |
| sales | annual sales (mill. dol.) |
| emply | employment (1000s) |
| invest | investment (mill. dol.) |
| rnd | R&D spending (mill. dol.) |
| cashfl | cash flow (= retained earnings + depreciation allowances) (mill. dol.) |
| kstock | knowledge stock (= accumulated R&D investment) (mill. dol.) |
| netcap | net capital stock (mill. dol.) |
| debt | long term debt (mill. dol.) |
| q | Tobin's q |
| loge | log (employment in 1000s) |
| rs | ratio of R&D invest to sales |
| cc | ratio of cashflow to net capital stock |
| drnd | dummy: zero R&D investment |
| exit | dummy: firm exits between year and year+1 |
| grsales | growth rate in sales (percent) between year and year+1 |

For this exercise, ignore the panel structure and treat the observations across years *as if* they were independent. Consider the Cobb-Douglas production function

$$\text{sales} = e^{\gamma}(\text{emply})^{\alpha}(\text{netcap})^{\beta}e^{\varepsilon},$$

where now the possible impact of knowledge stock on productivity is ignored, so that the model is log linear. You are concerned that emply may be contaminated by correlation with the disturbance.

a. Estimate the model in logs as a linear regression by OLS, using *only* the data for the years 1989-1994.

b. Estimate the model by instrumental variables, using capital stock variables and *lagged* emply as instruments.

c. Do an exogeneity test for emply.

d. If in step a, one had done OLS using data for *all* the years, how would you need to modify the exogeneity test procedure? (Hint: Look at what happens to the covariance matrix of the difference of the 2SLS and OLS estimates of the coefficients under the null.)