

PS207 Homework #2: Two Stage Models

For each problem, give me a couple of paragraphs writeup as well as the log file.

Problem 1: Download the data for the first problem by typing

```
use http://www.stata-press.com/data/r9/hsng2
```

This is US state level data from the 1980 Census. Suppose we wanted to estimate:

- (1) $hsngval = \alpha_0 + \alpha_1 faminc + \alpha_2 reg2 + \alpha_3 reg3 + \alpha_4 reg4 + \varepsilon$
- (2) $rent = \beta_0 + \beta_1 hsngval + \beta_2 pcturban + \nu$

where *hsngval* is the median dollar value of owner-occupied housing, *rent* is median monthly rent, *faminc* is median family income, and the *reg* variables indicate regions of the US.

Begin by running the appropriate models for equations 1 and 2, ignoring any possible endogeneity. Describe your results.

Now correct for any possible endogeneity in these equations using the `ivregress` command (`ivreg` in Stata 9 or earlier). See the online help for the syntax. Describe your results. Why might we expect to see endogeneity here? Were there any differences in these results compared to the estimates that ignored endogeneity?

Problem 2: Download the data for the second problem by typing

```
use http://www.stata-press.com/data/r9/laborsup
```

This is hypothetical data on 500 two person households. Suppose we wanted to estimate:

- (1) $fem_work = \alpha_0 + \alpha_1 fem_educ + \alpha_2 kids + \alpha_3 other_inc + \varepsilon$
- (2) $other_inc = \beta_0 + \beta_1 male_educ + \nu$

where *fem_work* is whether a woman is in the work force, *fem_educ* is that woman's level of education, *kids* is number of children, *other_inc* is non-work sources of income available to the woman, and *male_educ* is her male partner's level of education.

Begin by running the appropriate models for equations 1 and 2, ignoring any possible endogeneity. Describe your results.

Now correct for a possible endogeneity problem by using the `ivprobit` command. See the online help for the syntax. Describe your results. Were there any differences compared to the estimates that ignored endogeneity?