Problem Set 4. Instrumental Variables and Panel Data
(due Thursday, November 22, 2012)

1. Fertility and labor supply (ii)

In this problem set, we will work again with the data on married women from the US Census (available on the textbook website, www.pearsonhighered.com/stock_watson in the file **Fertility**, and described in the file **Fertility_Description**.) You already estimated OLS regressions relating weeks worked per year and number of children. The multiple regression analysis was unsatisfactory, so this week we will try IV.

a) The data set contains the variable *samesex*, which is equal to 1 if the first two children are of the same sex (boy-boy or girl-girl) and equal to 0 otherwise. Are couples whose first two children are of the same sex more likely to have a third child? Is the effect large? Is it statistically significant?

b) Explain if (and why) *samesex* is a valid instrument for the instrumental variable regression of *weeksworked* on *morekids*.

c) Is *samesex* a weak instrument?

d) Estimate the regression of *weeksworked* on *morekids* using *samesex* as an instrument. How large is the fertility effect on labor supply? Interpret magnitude and sign.

e) Is your instrument exogenous? Can you test it? If so, do it. If not, explain why not and what you would need in order to be able to do it.

f) Does the coefficient on *morekids* change a lot after instrumenting with *samesex*? (compared with an OLS regression without controls). Discuss and interpret the direction of the change.

g) Do the IV results change when you include the variables *agem1*, *black*, *hispan*, and *othrace* in the labor supply regression (treating these variables as controls)? Explain why or why not.

h) Can you test for the exogeneity of the instrument now?

i) Do you think you have now estimated the causal effect of fertility on labor supply credibly? Why or why not?

2. Capital punishment and crime (ii)

In Problem Set 1, you analyzed the association between capital punishment and crime in a cross-section of US States. Now we will extend the analysis to a panel data set.

a) You already have data on crime rates and capital punishment for all US states around 2010. Now, find information on the same two variables for two additional time periods (say, one year in the late 1980's and one in the late 1990's). On top of a binary variable indicating whether the death penalty is abolished or not in a state, you can also gather data on the annual number of executions. Create a Stata data set with the variables: crime rate, death penalty (binary), number of executions, state, and year. Show the table of summary statistics.

b) Compare the OLS results when using the three cross-sections separately.

c) Now, run a regression in first differences using the oldest and most recent periods only. Compare the results to the cross-sectional ones.

d) Which of the four regressions that you have estimated is your preferred one? Why?