Economics 250a -- Problem Set 3

In this problem set we will use data for married couples from the 2012 March CPS to examine married women's labor supply. We will also have a chance to look at the basic patterns in the "assortiveness" of couples' characteristics. To prepare for the problem set you should read:

Thomas A. Mroz. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." Econometrica, 55 (4) (July, 1987), pp. 765-799.

Mroz uses data for 753 married white women between the ages of 30 and 60, and fits models in the class we investigated in problem set #2. We will have around 4,000 observations, but will lack some of the covariates in his analysis, in particular information on parent's education. We do have information on immigrant and second generation status (variables imm and gen2).

The data set mroz.dta is on the course website, along with the program I used to assemble it from the March 2012 data (called extract2x.sas) and a do file, mroz2.do that you can use as a template. You can see how I made the data set by looking at extract2x. Since we want to use the wage measured in March (called "wage_ogr") as an alternative to the wage based on annual earnings divided by annual hours, I limit attention to married couples with (1) either H or W working last year and (2) either H or W working in March with a wage. This gives 4066 couples.

The variable labels should be fairly self-explanatory. I use prefix "h" for husband and "w" for wife. E.g.:
hwagesal = husband total earnings, which we will use as "nonwife income"
h_age = husband age; heduc = husband education; htwage=husband trimmed wage
wrgroup=category variable for wife race/ethnicity 1=wh 2=black 3=Hispanic 4=asian 5=other

1. Re-create Table III of Mroz as closely as you can with your data. Compare what happens using wlogwage (based on annual earnings / annual hours) and wlogwage_ogr based on wage in "OGR" (ie. reported in March for people who have a wage in March). Note that I defined a variable called "work_and_wage"=1 if a wife has hours last year and a wage in the OGR.

2. Look at Table IV and try to assemble as many of Mroz's basic instruments as you can.
Note: if you want to do some extra work, retrieve state unemployment rates for 2011 from the BLS web site. The variable "state" uses so-called FIPS codes for states, eg 93=California

3. To the extent possible, estimate the specifications 1-10 in Table IV. For each specification report the F-test for the first stage (an issue that was not well-understood at the time of Mroz's paper) and compute the implied compensated elasticity of labor supply. Try some of your own instruments for the women's wages.

4. To the extent possible reproduce the estimates in Table VII, specifications 1-4. For each specification report the F-tests for the two first stage equations and compute the implied compensated elasticity of labor supply. Try some of your own instruments for husband earnings and presence of children (eg. immigrant status predicts number of children).

5. Finally, we will try to estimate OLS and 2SLS models with a control function for participation (read section 1.5 carefully – we will only consider normal selection corrections). In the template program mroz2.do I show how to construct the correction term from a 1st stage probit model – this is the variable "lambda". Fit specifications 3 and 6 in Table IX.