**1)** Use a 50% random subsample of the wage-hours data attached to your problem set (MOM data[1], variables are respectively `lnhr lnwg kids ageh agesq disab id year`) and estimate $\beta$ as below:

$$Inhr_{it} = a_i + \beta\, lnwg_{it} + E_{it},$$

a) Can $\beta$ be directly interpreted as a labor supply elasticity? Explain.

b) For the following estimators: (1) pooled OLS, (2) between, (3) within, (4) first differences, (5) random effects GLS, (6) random effects MLE give (i) $\beta$ (estimated coefficient of Inwg), (ii) default standard error, and (iii) panel bootstrap standard error with 200 replications.[2]

c) Are the estimates of $\beta$ similar?

d) Is there a systematic difference between default standard errors and panel robust standard errors?

e) Will the pooled OLS estimator in part (b) be consistent for $\beta$ in a fixed effects model? Will the pooled OLS estimator be consistent for $\beta$ in a random effects model?

f) Perform a Hausman test of the difference between the fixed and random effects (GLS) estimates of in this model. Do this manually using the earlier regression output with the default standard errors. What do you conclude and which model is favored?

g) Given the preceding evidence, do you believe that the labor supply curve is upward sloping? Explain.

**2)** Use the data in NORWAY.RAW[3] for the years 1972 and 1978 for a two-year panel data analysis. The model is a simple distributed lag model:

$$log(crime_{it}) = \theta_0 + \theta_1\, d78_t + \beta_1\, clrprc_{i,t-1} + \beta_2 clrprc_{i,t-2} + C_i + u_{it}.$$

The variable clrprc is the clear-up percentage (the percentage of crimes solved). The data are stored for two years, with the needed lags given as variables for each year.

a) First estimate this equation using a pooled OLS analysis. Comment on the deterrent effect of the clear up percentage, including interpreting the size of the coefficients. Test for serial correlation in the composite error $v_{it}$ assuming strict exogeneity

[1] This dataset is based on J. Ziliak (1997)

[2] MLE random effects and bootstrap standard error are optional and have extra credit.

[3] Use http://www.stata.com/data/jwooldridge/eacsap/norway

b) Estimate the equation by FE, and compare the estimates with the pooled OLS estimates. Is there any reason to test for serial correlation? Obtain heteroskedasticity robust standard errors for the FE estimates.

c) Using FE analysis, test the hypothesis Ho : $\beta_1 = \beta_2$. What do you conclude? If the hypothesis is not rejected, what would be a more parsimonious model? Estimate this model.

**3)** The data in WAGEPAN.RAW[4] are from Vella and Verbeek (1998) for 545 men who worked every year from 1980 to 1987. Consider the wage equation:

$$log(wage_{it}) = \theta_t + B_1 educ_i + \beta_2 black_i + \beta_3 hisp_i + \beta_4 exper_{it}$$

$$+ \beta_5 exper_{it}^2 + \beta_6 married_{it} + \beta_7 union_{it} + C_i + u_{it}.$$

The variables are described in the data set. Notice that education does not change over time.

a) Estimate this equation by pooled OLS, and report the results in standard form. Are the usual OLS standard errors reliable, even if $c_i$ is uncorrelated with all explanatory variables? Explain. Compute appropriate standard errors.

b) Estimate the wage equation by RE. Compare your estimates with the pooled OLS estimates.

c) Now estimate the equation by FE. Why is $exper_{it}$ redundant in the model even though it changes over time? What happens to the marriage and union premiums as compared with the RE estimates?

d) Now add interactions of the form d81.educ, d82.educ, ... ,d87.educ and estimate the equation by FE. Has the return to education increased over time?

e) Return to the original model estimated by FE in part c. Add a lead of the union variable, $union_{i,t+1}$ to the equation, and estimate the model by FE (note that you lose the data for 1987). Is $union_{i,t+1}$ significant? What does your finding say about strict exogeneity of union membership?

[4] Use http://www.stata.com/data/jwooldridge/eacsap/wagepan