Homework 8 (3 questions). Due Friday November 5.

1. Use the posted bootstrap program to estimate a linear regression model.
1) Try samples N=10, 20, 50, 100 and run the program a few times. In this model, at what sample size does it look as if the bootstrap estimator of standard errors of the slope parameters is OK (similar to the parametric estimator).
2) Try and vary the number of bootstrap replications. How many replications seems to be needed for the estimator to work?
3) Try and simulate the model with error terms that are Cauchy distributed (use the ratio of standard normals). Those are pretty crazy and the parametric standard errors may be off. Keep N=100 and do a small Monte Carlo study (maybe 50 iterations, start with a low number and see what your computer can handle in not-too-long time) and report: a) the average parametric standard errors; b) the average bootstrap standard errors; c) the standard error of the estimated parameters across you Monte Carlo iterations (that is the true standard error, at least if the number of iterations is not too low). Is the bootstrap standard errors of the parametric OLS standard errors better for the model with Cauchy errors?

2. Verify formula (3) in Moulton's article for the simple case of m=3. Assume the matrix of regressors is

$$X = \begin{pmatrix} x_1 \\ x_1 \\ x_1 \\ x_2 \\ x_2 \\ x_2 \end{pmatrix},$$

and the error variance matrix is

$$V = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & 0 & 0 & 0 \\ \rho & 1 & \rho & 0 & 0 & 0 \\ \rho & \rho & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \rho & \rho \\ 0 & 0 & 0 & \rho & 1 & \rho \\ 0 & 0 & 0 & \rho & \rho & 1 \end{pmatrix}.$$

3. a) Use the updated Matlab panel data program posted and run it without fixed effects, but with clustered standard errors, clustered by state, and report the t-statistics calculated from the program. (Note, that you are not to use any of the built-in Matlab commands for this.)

b) Simplify the code to calculate White standard errors, then run the program, report the t-statistics and compare to a) .

c) Include state fixed effects and repeat questions a) and b).

d) Now assume that the errors are clustered by time (year), rather than by state. Modify the program and run it again. Report the t-statistics. Do this with or without time fixed effects.