

**Homework 8. Due Wednesday November 2nd.**

1. Use the program `bootstr.txt` 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. Use a CPS Stata dataset from Cameron's web-page. Run a regression. (Wages on some of the other variables.) Compare the significance of the parameters using standard OLS standard errors, White standard errors, and clustered standard errors? (I haven't tried this out myself. IT is not obvious that clustering is important if you don't use aggregate regressors or panel data, so try one of these features, like a dummy for the state.)

3. Verify that the integral (8.43) in Davidson and MacKinnon (p. 327) doesn't have a finite solution.