CODERS' CORNER

WILD-BOOTSTRAPPED T-STATISTICS

A **small number of clusters** seriously **affects the validity of clustered standard error.**

Wild-bootstrapped t-statistics perform well, even with numbers of clusters as low as 10 (Cameron et al. 2008). As a useful rule of thumb, you should **use wild bootstraps** for inference when your **number of clusters in at least one clustering dimension is below 40**.

Wild bootstraps are implemented in Stata with the **boottest** command.

The command saves the p-value of the tested hypothesis in r(p) and plots the empirical distribution of the test-statistic.   .

The syntax:
First run your normal regression (e.g. `reg y treatment`). Then you can use the following syntax:

```
boottest treatment=0, cluster(cluster1 cluster2) reps()  seed()
```

- `treatment=0` is the hypothesis of interest. You can also specify more than one hypothesis and bootstrap Wald-tests of significance.
- `cluster(cluster1 cluster2)` defines the clustering dimensions. The boottest command allows multiple dimensions of clustering and bootstraps using all specified clustering dimensions.
- `reps` specifies the number of bootstrap replications.
- `seed` is a random seed for the boottest command. You should always set a seed to ensure that others can replicate your results.

Lukas Hensel, St Peter's College, Oxford
25 February 2019