Advanced Econometrics II: 2023
Assignment on Panel Data
Deadline: January 31 at 5pm, 2023

---

1. Monte Carlo Study

In this assignment you investigate the properties of the FE and the AH estimators (with instruments $y_{i,t-2}$ and $y_{i,t-3}$ in levels)[1] in the stylized AR(1) model:

$$y_{i,t} = \alpha_0 y_{i,t-1} + \eta_i + \varepsilon_{i,t}, \quad t = 1, \ldots, T, \tag{1}$$

where $\varepsilon_{i,t} \sim N(0,1)$ while $y_{i,0}$ is drawn from a semi-stationary distribution:

$$y_{i,0} = \delta \frac{\eta_i}{1-\alpha_0} + \varepsilon_{i,0}, \quad \varepsilon_{i,0} \sim N(0, (1-\alpha_0^2)). \tag{2}$$

Finally, the individual effect is $\eta_i \sim N(0, \sigma_\eta^2)$. All stochastic quantities above are iid over all (i,t).

In what follows you fix the number of replications to $M = 4000$, the true value $\alpha_0 = 0.5$, and the nominal size for $t-$test at 5%. Consider the following parameter space:

$$N = \{100; 150\}, \quad T = \{6; 15\}, \quad \delta = \{0.5; 1\}, \quad \sigma_\eta^2 = \{1; 2\}. \tag{3}$$

Thus there are in total 16 different combinations of $(N, T, \delta, \sigma_\eta^2)$ that you need to consider.

Remark 1. For simplicity, you can consider only the one-step GMM estimator for the AH moment conditions with the *homoscedasticity-only* weighting matrix. However, do not forget to use the appropriate formulae to calculate the variance-covariance matrix that should be consistent even if the true errors are heteroscedastic. Also use the heteroscedasticity robust variance-covariance matrix estimator for the FE estimator.

[1] Hence, the degree of overidentification is 1.

January 26, 2023

Analysis

1. Report the Monte Carlo bias and the Root Mean Squared Error (RMSE) for both estimators.
2. Report empirical rejection frequencies of the tests based on the corresponding t-statistics for the null hypothesis of $H_0 : \alpha = \alpha_0$.
3. Discuss extensively these results.
4. Do you observe any patterns in terms of the relative dominance of the AH estimator over the FE estimator?
5. Consider the Half Panel Jackknife bias-corrected FE estimator. Does this approach provide an improvement over the FE estimator in terms of the bias and/or RMSE?
6. Report empirical rejection frequencies for the HPJ estimator based on the S.E. of the FE estimator. Comment on the results.

Finally, repeat your analysis using the same DGP but this time will all your estimators being defined in terms of cross-sectionally de-meaned data, i.e. subtract cross-sectional averages from your data (for each variable and all time periods) before conducting analysis. Do your conclusions change when using de-meaned data vs. the original data? Comment.

Comments on the Implementation

1. It is convenient to use the Python code provided as a basis for your Monte Carlo study. It contains a function to generate artificial data (*Generate Data*) as well as the skeleton on how to setup the MC study.
2. Avoid loops within the functions for estimators as much as possible. On the other hand, loops are difficult to avoid in the construction of the Standard Errors.
3. Note that for the estimators on de-meaned data you do not have to re-run your Monte Carlo analysis. You simply need to consider more estimators (6 in total) for each design.

Submission

1. Submit your report which should not exceed 5 pages. Make sure that you discuss the results and draw some conclusions. You can also use figures yo make your conclusions stronger.
2. Submit your Python codes and/or Notebooks. Your code should be well written and understandable.
3. It goes without saying, that I should be able to run the code without any errors.
4. Explicitly mention the packages (beyond NumPy or Pandas) that are needed to run your codes.