ReadMe file for the replication package accompanying "Heterogeneous Autoregressions in Short T Panel Data" by M. Hashem Pesaran and Liying Yang

June 22, 2024

- The zipped file "MC_replication.py.jae.2024.zip" contains R and MATLAB codes and results for the Monte Carlo simulation in Section 7 of the main paper and Section S.8 of the online supplement.

- The zipped file "PSID_application.py.jae.2024.zip" contains data files and associated R and MATLAB codes for the empirical application (earnings dynamics) in Section 8 of the main paper and Section S.9 of the online supplement.

- Each of these files contains its own ReadMe files.

MONTE CARLO SIMULATIONS
========================

Estimators
==========

"FDAC" refers to the proposed estimator based on the autocorrelations of first differences, and "HetroGMM" refers to the proposed generalized method of moments (GMM) estimator based on the autocovariances of first differences in the paper. See Section 6 of the main paper.

"HomoGMM" refers to several GMM estimators proposed in the literature for homogenous autoregressive (AR) panels, including the estimators proposed by Anderson and Hsiao (1981, 1982) (AH), Arellano and Bond (1991) (AB), Blundell and Bond (1998) (BB), and the augmented Anderson-Hsiao (AAH) estimator proposed by Chudik and Pesaran (2021), as well as the FDLS estimator due to Han and Phillips (2010). See Sections 4 and 7.1 of the main paper for details.

"MSW" refers to a kernel-weighted likelihood estimator proposed by Mavroeidis et al. (2015) for heterogeneous AR panels. See Section 1 of the main paper for a brief review.

Programs
========

1. The MC experiments comparing the proposed FDAC and HetroGMM estimators
----

-- "HP_AR_MC_FDAC_m1.R" simulates and saves the MC results comparing FDAC and HetroGMM estimators of the mean of heterogeneous AR coefficients.

-- "HP_AR_MC_FDAC_m1_exp_tab_fig.R" generates Tables S.1, S.2, S.5, S.12 and S.13, and Figures S.1, S.4 and S.5 in Section S.8 of the online supplement. The tables will be saved in "HetroAR_MC_FDAC_moments_exp_m1.xlsx".

-- "HP_AR_MC_FDAC_var.R" simulates and saves the MC results comparing FDAC and HetroGMM estimators of the variance of heterogeneous AR coefficients.

-- "HP_AR_MC_FDAC_var_exp_tab_fig.R" generates Tables S.3, S.4, S.6 and S.7, and Figure S.2 in Section S.8 of the online supplement. The tables will be saved in "HetroAR_MC_FDAC_moments_exp_var.xlsx".

2. The MC experiments comparing FDAC and HomoGMM estimators

All relevant simulation codes for this set of MC simulations are stored in the subfolder "MC_replication.py.jae.2024/FDAC vs HomoGMM".

-- "HP_AR_MC_FDAC_vs_HomoGMM_exp.m" simulates and saves the MC results comparing FDAC and HomoGMM estimators of homogeneous and the mean of heterogeneous AR coefficients.

-- "HP_AR_MC_FDAC_vs_HomoGMM_tab_fig.R" generates Tables S.8, S.9, S.10, S.14, S.15 and S.16 and Figure S.3 in Section S.8 of the online supplement. The tables will be saved in "HetroAR_MC_FDAC_vs_HomoGMM_results.xlsx".

3. The MC experiments comparing FDAC and MSW estimators

-- "HP_AR_MC_FDAC_vs_MSW_exp.R" simulates and saves the MC results comparing FDAC and MSW estimators of the mean of heterogeneous AR coefficients.

-- "HP_AR_MC_FDAC_vs_MSW_exp_tab.R" generates Tables S.11 and S.17 in Section S.8 of the online supplement. The tables will be saved in "HetroAR_MC_FDAC_vs_MSW_results.xlsx".

Simulation results
==================

-- The subfolder "MC_replication.py.jae.2024/Simulation results/FDAC vs HetroGMM" contains simulation results comparing FDAC and HetroGMM estimators shown in Tables S.1-S.7 and S.12-S.13 and Figures S.1-S.2 and S.4-S.5.

-- The subfolder "MC_replication.py.jae.2024/Simulation results/FDAC vs HomoGMM" contains simulation results comparing FDAC and HomoGMM estimators shown in Tables S.8-S.10 and S.14-S.16 and Figure S.3.

-- The subfolder "MC_replication.py.jae.2024/Simulation results/FDAC vs MSW" contains simulation results comparing FDAC and MSW estimators shown in Tables S.11 and S.17.

EMPIRICAL APPLICATION
=======================

Data
====

-- "PSID2.csv" contains a sample of households drawn from the Panel Study of Income Dynamics (PSID) data for the years 1976-1995, retrieved from the website: https://psidonline.isr.umich.edu. See "PSID2 1967-1996 data and notes.xlsx" for variable descriptions, sample selection criteria, and summary statistics.

-- "PCED2.csv" contains the GNP personal consumption expenditure deflator (using 1996 as the base year), retrieved from the Bureau of Economic Analysis of the US Department of Commerce in 2021.

-- The following data files contain the five- and ten-year sub-panels of the deflated PSID data for computing MSW estimates.
- "PSID2_educ_76_80.mat"
- "PSID2_educ_76_85.mat"
- "PSID2_educ_81_85.mat"
- "PSID2_educ_81_90.mat"
- "PSID2_educ_86_90.mat"
- "PSID2_educ_86_95.mat"
- "PSID2_educ_91_95.mat"

Programs
========

-- "example_ar1.m" loads data from "PSID2.csv" and "PCED2.csv", computes AAH, AB, BB, and FDAC estimates, and saves the estimation results in "PSID2_educ.mat".

-- "Hetro_AR_PSID_msw_parallel.R" loads functions stored in "Hetro_AR_PSID_msw_functions.R" for computing the MSW estimator and loads data listed above, and saves the estimation results in "Hetro_AR_PSID_msw.RData".

-- "Hetro_AR_PSID_tab.R" loads estimation results from "PSID2_educ.mat" and "Hetro_AR_PSID_msw.RData" and generates Table 1 in Section 8 of the main paper and Tables S.19-S.23 in Section S.9 of the online supplement. The tables will be saved in "HetroAR_PSID_educ_results.xlsx".

Custom data sets
================

For estimation of panel AR(1) model using a custom data set please use "FDAC_custom_data_estimation.R". Please modify the code in line 12 of the R script to ensure that the custom data set is correctly input into the estimation function. The estimation results will be saved in "FDAC_results.xlsx".

===
- The MATLAB scripts were written and executed using MATLAB version R2022b.
- The R scripts were written and executed using Rstudio version 2023.12.0+369.
- Please ensure that the path is correctly set in both MATLAB and R scripts to the directory containing the data files and the results of the simulations and empirical estimation, and the packages loaded in the R scripts are installed on either local desktops or clusters.
===

We are grateful for the helpful comments and suggestions offered by Hayun Song.

References

Anderson, T. W., and Hsiao, C. (1981). Estimation of dynamic models with error components. Journal of the American Statistical Association 76, 598-606.

Anderson, T. W., and Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. Journal of Econometrics 18, 47-82.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies 58, 277-297.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115-143.

Chudik, A., and Pesaran, M. H. (2022). An augmented Anderson–Hsiao estimator for dynamic short-T panels. Econometric Reviews, 41, 416-447.

Han, C. and P. C. Phillips (2010). GMM estimation for dynamic panels with fixed effects and strong instruments at unity. Econometric Theory 26, 119-151.

Mavroeidis, S., Sasaki, Y., and Welch, I. (2015). Estimation of heterogeneous autoregressive parameters with short panel data. Journal of Econometrics, 188, 219-235.