Replication files for "Robust Inference in Models Identified via Heteroskedasticity"

Daniel Lewis

July 22, 2020

This folder contains the Matlab code needed to replicate the empirical analysis and Monte Carlo study of the paper "Robust Inference in Models Identified via Heteroskedasticity".

- The data needed for replication can be freely downloaded from the replication archive for Nakamura and Steinsson (2018), here. I do not include the raw files in this replication folder directly. In particular, to run the empirical analysis (starting with `Compile_dataset.m`), the files "`RealYields.csv`", "`NominalYields.csv`", and "`dataforMatlabAltEstTick.csv`" from the replication archive should be added to the Matlab path. Alternatively, I provide the compiled dataset "`dataset.mat`" for direct use in "`Empirical_application.m`".
- To run the Monte Carlo study and produce the figures, the functions "`nearestSPD.m`" (with light modification) and "`subplot1.m`" should be downloaded from Mathworks and added to the Matlab path. See details below.
- The code was run in Matlab 2019b on a 2.5GHz MacBook Pro laptop using Mojave 10.14.6 OS.

1 Data

1. `RealYields.csv`: this file can be found in the replication archive for Nakamura and Steinsson (2018). Contains real 2-year forward rate as `TIPSF02`.
2. `NominalYields.csv`: this file can be found in the replication archive for Nakamura and Steinsson (2018). Contains nominal 2-year Treasury yield and 2-year forward rate as `SVENY02` and `SVENF02`.

3. `dataforMatlabAltEstTick.csv`: this file can be found in the replication archive for Nakamura and Steinsson (2018). Contains 30-min window policy news shocks and FOMC announcement indicator as `path_intra_wide` and `FOMCused`.

4. `Compile_dataset.m`: this Matlab file reads in data from `RealYields.csv`, `NominalYields.csv`, and `dataforMatlabAltEstTick.csv`, transforms each data series as necessary (e.g., daily changes), converts each data series to timetable form, synchronizes the timetables, and saves the resulting full-sample, control regime, and policy regime data series as `dataset.mat`. The sample period is defined as those policy dates and control dates from the Nakamura and Steinsson (2018) policy news shock dataset for which all variables are available (4 January 2004 to 19 March 2014, in practice), excluding the July 2008-June 2009, as in the original paper. The variables output are:

    - `RFc`, `RFp`, `RF`: 1-day changes in 2-year real forward rates on control days, policy days, and all days (stacking control and then policy), respectively.
    - `NFc`, `NFp`, `NF`: 1-day changes in 2-year nominal forward rates on control days, policy days, and all days (stacking control and then policy), respectively.
    - `NYc`, `NYp`, NY: 1-day changes in 2-year nominal Treasury yields on control days, policy days, and all days (stacking control and then policy), respectively.
    - `HFc`, `HFp`, `HF`: 30-minute "policy news shocks" from Nakamura and Steinsson (2018) on control days, policy days, and all days (stacking control and then policy), respectively.

- `dataset.mat`: intermediate compiled dataset ready to be loaded by `Empirical_application.m`.

2 Empirical analysis

1. `Empirical_application.m`: this file loads the empirical dataset, `dataset.m`, conducts $F-$statistic based tests for weak identification, and estimates the four specifications reported in the paper. It computes both standard and non-robust confidence sets for each parameter in H, and thus also provides results for the Andrews (2018) test for weak identification. It compiles these results into Tables 4 and 5 of the paper.

3 Monte Carlo study

1. `size_MC.m`: this file conducts the Monte Carlo study assessing the size of standard inference procedures, projection tests, and the subset test proposed in the paper for an empirically calibrated model, Tables 2 and 3 in the paper. It also produces Figure 2 based on the output and saves quantiles of the test statistics under the null hypothesis, `quanttab.mat`, to compute size-adjusted power (SAP).

2. `power_MC.m`: this file conducts the Monte Carlo study assessing the power of standard inference procedures, projection tests, and the subset test proposed in the paper for an empirically calibrated model, Figures 3-5 in the paper. It uses the quantiles computed by `size_MC.m`, `quanttab.mat`.

4 Auxiliary functions

1. `Figure1.m`: this file produces Figure 1 in the paper by simulating data from DGPs with parameters chosen aribtrarily to illustrate the weak identification problem.

2. `WeakHetFirstF.m`: this file is not used in the paper, but is supplied for use by empirical researchers wishing to conduct the heteroskedasticity-robust first-stage $F-$statistic for the 2-dimensional "simple case" model where only the variance of the shock of interest changes. It takes as inputs `eta21`, $T_1 \times 1$ vector of regime 1 observations of the innovation that is normalized to have unit loading on the shock of interest, and `eta22`, $T_2 \times 1$ vector of regime 2 observations of the same innovation. It is assumed that the variance of the shock of interest changes across regimes, while the variance of the other shock does not. In the paper, the inputs correspond to η_{2t} in the control regime and η_{2t} in the policy regime, respectively. The function outputs `Fstat`, the $F-$statistic from the first-stage regression, and optionally `reject`, indicators for rejection of the null of weak identification against the critical values from Montiel Olea and Pflueger (2013) at the 0.05, 0.1, 0.2, and 0.3 bias levels.

3. `InvertTest.m`: this function computes robust confidence sets using the subset inference approach based on Theorem 3. It inverts the test statistic over a grid of values for the parameter of interest. It takes as inputs `theta`, the estimated parameter vector, `etasq1`, the $T_1 \times 3$ unique vectorization of the outer product of innovations in regime 1; `etasq2`, the $T_2 \times 3$ unique vectorization of the outer product of outer product of innovations in regime 2; `pind`, an index for the parameter of interest; `range`, a grid over which to perturb the parameter of interest; and `level`, the level of the

test. It outputs `CIbnds`, a robust confidence interval for the parameter of interest, and optionally `DistCI`, confidence intervals under a set of relative distortions for an Andrews (2018) identification test and `stats`, the full grid of test statistics.

4. `momfast.m`: for a 2-variable model, this function computes the GMM objective function based on equations (3) and (4). It takes as inputs `etasq1`, the $T_1 \times 3$ unique vectorization of the outer product of innovations in regime 1; `etasq2`, the $T_2 \times 3$ unique vectorization of the outer product of outer product of innovations in regime 2; `T`, the total sample length; and `p`, the parameter vector. It returns `obj`, the GMM objective function using the efficient weighting matrix; `f_T`, the sum of moments; `omega`, the covariance of moments; and `matinv`, the efficient weighting matrix.

5. `momfastsub.m`: for a 2-variable model, this function computes the GMM objective function based on equations (3) and (4), conditional on a fixed value for some parameter of interest in H. `etasq1`, the $T_1 \times 3$ unique vectorization of the outer product of innovations in regime 1; `etasq2`, the $T_2 \times 3$ unique vectorization of the outer product of outer product of innovations in regime 2; `T`, the total sample length; `nuis`, the vector of nuisance parameters; `Hint`, the value for the parameter of interest in H; and `pind`, the indicator for which free parameter in H is the parameter of interest (1 for H_{21}, 2 for H_{12}). It returns `obj`, the GMM objective function using the efficient weighting matrix; `f_T`, the sum of moments; `matinv`, the efficient weighting matrix; and `omega`, the covariance of moments.

6. `nearestSPD.m`: this function is John d'Errico's function to compute the nearest symmetric positive definite matrix to an input matrix `A`. It is not included in the replication file, but is freely available for download from Mathworks File Exchange, (https://www.mathworks.com/matlabcentral/fileexchange/42885-nearestspd). I use a version slightly modified to throw an error if the matrix `A` is still not SPD after 1000 iterations. In particular, I add

   ```
   if k==1000
      error('Too many iterations')
   end
   ```

   at line 59.

7. `subplot1.m`: this function is Eran Ofek's function to provide greater control over subplot behaviour. It is not included in the replication file, but is freely available for download from Mathworks File Exchange,

(https://www.mathworks.com/matlabcentral/fileexchange/9694-an-improved-subplot-function).

References

D'Errico, John, 2013. "nearestSPD", https://www.mathworks.com/matlabcentral/fileexchange/42885-nearestspd, MATLAB Central File Exchange.

Nakamura, Emi and Steinsson, Jon, 2018, "Replication Data for: 'High-Frequency Identification of Monetary Non-Neutrality: The Information Effect'", https://doi.org/10.7910/DVN/HZOXKN, Harvard Dataverse, V2.

Ofek, Eran, 2006. "An improved subplot function", https://www.mathworks.com/matlabcentral/fileexchange/9694-an-improved-subplot-function MATLAB Central File Exchange.