

README

Paper: Decentralization Estimators for Instrumental Variable Quantile Regression Models

Authors: Hiroaki Kaido and Kaspar Wuthrich

The replication package includes the following files.

- *algorithms_final.R*
Collection of generic functions used by the other R scripts. The functions require the installation of R-packages.
- *401k_empirical_application_final.R*
R script to replicate the results for the empirical application.
- *sim_dgp401k_ct_final.R*
R script to replicate the comparison of the computation times with 1 and 2 endogenous regressors.
- *sim_dgp401k_inference_final.R*
R script to replicate the simulation results on the coverage of the confidence intervals based on the empirical application. This file is designed to be executed from the command line. See the next section for details.
- *sim_dgp401k_bias_rmse_final.R*
R script to replicate the simulations evaluating bias and RMSE with 1 and 2 endogenous regressors based on the empirical application.
- *sim_dgp401k_3endog_final.R*
Simulation study to compute bias, RMSE, and average computation time with 3 endogenous regressors based on the empirical application.
- *sim_locscal_final.R*
R script to replicate the simulation results on the bias and RMSE based on the location-scale DGP.
- *sim_locscal_inference_final.R*
R script to replicate the simulation results on the coverage properties of the confidence intervals based on the location-scale DGP. This file is designed to be executed from the command line. See the next section for details.
- *restatw.dat*
Dataset for the empirical application. These data have been previously analyzed by Chernozhukov and Hansen (2004, Review of Economics and Statistics) and were obtained from Chris Hansen.

The simulations results are saved in the folder RData. The figures are saved in the folder Figures.

The simulation results on the coverage of the confidence intervals were obtained from the cluster at Boston University based on the R-code provided in the replication package.

Instructions:

Running most of the R files should be straightforward. First, install the packages used by *algorithms_final.R* and the R file you are using for replication. In each R file, there is a line that specifies the working directory using the R function “setwd”. Replace its argument with your working directory. Running the R file will generate and save results of a Monte Carlo simulation or an empirical application.

There are two files that are designed to be executed from the command line. They are *sim_dgp401k_inference_final.R* and *sim_locscal_inference_final.R*. They can be executed from the command prompt using the following syntax.

```
Rscript sim_dgp401k_inference_final.R [# of cores] [ind]
```

```
Rscript sim_locscal_inference_final.R [# of cores] [ind]
```

For each command, the first argument (# of cores) should be set to the number of CPU cores (or threads) you are willing to use for parallelization. The second argument “ind” is used to determine the combination of the nominal level (alpha) and quantile index (tau) to be used. The following table summarizes the correspondence.

		tau				
		0.15	0.25	0.5	0.75	0.85
alpha	0.1	ind = 1	ind = 2	ind = 3	ind = 4	ind = 5
	0.05	ind = 6	ind = 7	ind = 8	ind = 9	ind = 10

For example, if you want to replicate the simulation result based on the 401k example with alpha = 0.05 and tau = 0.5 using 4 CPUs, you can use the following command.

```
Rscript sim_dgp401k_inference_final.R 4 8
```