

Code Documentation for Aristodemou and Rosen (2022) "A Discrete Choice Model for Partially Ordered Alternatives"

This readme file provides a description of the files contained in AristodemouRosen_2022_QE.zip used to obtain the tables and figures in the paper. The top level of the zip file is divided into two directories, named "Monte Carlos" and "Application", that contain code for simulation results and the empirical application, respectively. The code for the application requires access to the data source.

The application folder contains files with Stata and R code. The Stata do file in this folder creates the data extracts that are then used by the R code and must be run first.

Data source

The purchase data are from the Kantar FMCG Purchase Panel. We use data from 2004-2005. These data are available commercially. The data are also available at the IFS for the use of academic visitors after signing a confidentiality agreement.

Monte Carlos Folder

Files "MC_Test_Script DGP1.R", "MC_Test_Script DGP2.R", "MC_Test_Script DGP3.R" produce the Monte Carlo results reported in Tables 16-18 in Appendix D. They use functions defined in "POR_Likelihood_Functions_Final.R", "POR_MC_Functions.R" and "POR_Sim_Linear.R".

"Compute_POR_Elasticities.R" produces elasticity estimates shown in Table 1 of the online supplement, obtained from simulation data from DGP2, as described in the online supplement.

Application Folder

The file "Estimation_Data_Creation_Aristodemou_and_Rosen_2022_QE.do" accesses the data and creates output for Tables in the paper. In addition, it creates two data extracts that are subsequently used by the R files also in this folder. The Stata do file needs to be run first, prior to running the R code.

The file "Compute_Tables_4_5.R" contains files that produce the results reported in Tables 4 and 5.

The default is to use the data that corresponds to that in Table 4. To change to that used for Table 5, change the variable "dtapath" as indicated in lines 7 and 8.

The default setting is also to compute point estimates for theta, Wald confidence intervals for theta, and QLR confidence intervals for only the first four components of theta. To produce QLR confidence intervals for other components of theta, change the variable parameter_indices in line 56 as indicated in the comment. In our experience it took roughly one to four days per parameter component to run. We computed these in batches of four at a time, executing the entire file with each of the following separately:

```
parameter_indices <- seq(1,4,1)
```

```
parameter_indices <- seq(5,8,1)
```

```
parameter_indices <- seq(9,12,1)
```

```
parameter_indices <- seq(13,16,1)
```

The files "Compute_POR_Elasticities_Table_6.R" computes elasticities and creates the output used in Table 6, as described in the paper.

The files "POR_Likelihood_Functions_Final.R" and "POR_Load_Data_Functions.R" contain functions that are used by the other R files described above.