**Code for “Consumption Heterogeneity: Micro Drivers and Macro Implications”**

**Empirical Code**

To replicate the results and graphs for the paper you need to take the following steps, first on the Danmarks Nationalbank research server, and then on your local machine.

On the **research server** at Statistics Denmark:

1. Run "Merge family and firms.sas" in the "SeverSAScode" folder in SAS. The program exports a .txt-file. This is very slow (overnight).
2. Run "Read data.do" in the "SeverStatacode" folder in Stata. The imports the .txt file and creates a Stata dataset from it. You may need to change the path from which the file reads the .txt file to match the output from step 1.
3. Run "Create\_residuals\_CALC\_DECILES.do" in the "SeverStatacode" folder in Stata.
   1. Make sure the paths to rawdata, savedirectory, dofiles and logfiles are all correct
   2. Set level=1 to calculate residuals in levels, level=0 to calculate in logs.
   3. Set labor\_inc\_only=1 to use labor income, labor\_inc\_only=0 to use total income
   4. You can also save time by setting five\_percent\_sample=1 (code will be run on a 5% sample), and you can label the log files using the global "run".

This Stata do file produces a file containing income and consumption residuals, after controlling for observables.

1. Run "Descriptives.do" in the "SeverStatacode" folder in Stata
   1. Make sure the paths to rawdata, savedirectory, dofiles and logfiles are all correct

This outputs descriptive statistics for the full population as well as for the estimation sample to a log file.

1. Run "Create\_csvfile\_for\_R.do" in the "SeverStatacode" folder. This creates the income consumption residuals file in a format ready to run in R.
2. Run "Create\_and\_save\_moments.R" in the "SeverRcode" folder in R.
   1. Set Booleans five\_percent\_sample, levels, lincome\_head and lincome\_spouse appropriately. lincome\_head and lincome\_spouse indicate whether to use the labor income of the head, spouse or total income (if both set to 0)
   2. Ensure Rcode\_folder, empirical\_input\_folder, moments\_dir and log\_dir all point to the correct directories.

This file will create the empirical moments that are then used in estimation. These moments are stored in the "moment\_dir" directory.

1. Copy the the relevant moments from the "moment\_dir" directory on the server and store them locally

On your **local** machine:

1. Run "MAIN.R" in the "Rcode" folder in R.
   1. Ensure Rcode\_folder, moments\_dir, figures\_fir, tables\_dir and PythonResults\_folder all point to the correct folders. PythonResults\_folder should contain the results of the modeling exercise (documented later in this help file)
   2. Choose "tag" appropriately to reflect whether to use level/log labor/total, head/spouse data.

This produces most of the figures in the paper. You may get a message saying “Too many open devices” in which case close some of the graph windows that have been produced and run the remaining code.

1. Run "Auclert\_details\_plot.R" in the "Rcode" folder in R. This produces the distribution of MPX by income, NNP and URE, adding in extra information such as homeownership.
2. Run "Robustness\_simulations.R" and "time\_agg\_random\_walk\_graph.R" in the "Rcode" folder in R to produce some of the robustness results and graphs on time aggregation that appear in the paper.

Robustness checks that are not implemented in the code can be run by repeating the server steps above, but changing the sample in the Stata do file appropriately for the robustness check, and storing data in separate folders.

**Model Code**

The model code can be found in the "PrefShockModel" folder, along with code that prints some of the tables in the paper. All this code is Python code and requires the HARK toolkit to run. See <https://econ-ark.org/> for more details

1. Run "MAIN.py" in Spyder.
   1. Set Booleans estimate\_benchmark and estimate\_pref\_shock to determine if you want to estimate these models, or use the estimated numbers hard-coded in the file. Note estimation is very slow.

This runs both the benchmark model and the preference shock model and stores the results for use in R.

1. Run " PrintSummaryStats.py" in Spyder in order to create the summary tables and URE tables in the paper.