

The replication package includes the following files

File name	Sub-directory	Type
Data files		
cpi_detailed.dta		STATA data file
init_cond.dta		STATA data file
min_wage.dta		STATA data file
poverty_line2.dta		STATA data file
price_indices.dta		STATA data file
psid_states.dta		STATA data file
table_2_3_5_annual.csv		CSV file
data4estimation.dta	output	STATA data file
psid_cons.dta	output	STATA data file
STATA programs (by order of appearance in the code)		
shell.do		STATA do file
psid_interview.do		STATA do file
psid_individual.do		STATA do file
psid_consumption.do		STATA do file
psid_sample.do		STATA do file
prepare_bootstrap.do		STATA do file
residual_measures.do		STATA do file
advanced_information_test.do		STATA do file
bootstrap.do		STATA do file
variance_estimation.do		STATA do file
tax_rate_estimation_snap.do		STATA do file
pi_s_estimation.do		STATA do file
frisch_estimation.do		STATA do file
aggregate_consumption.do		STATA do file
descriptive_stats.do		STATA do file
tables_7_8_9.do		STATA do file
pi_s_by_age_figure.do		STATA do file
calculate_AT_elasticities.do		STATA do file
marshallian_figure.do		STATA do file
moment_fit_table.do		STATA do file
external_fit.do		STATA do file
identification_figures.do		STATA do file
identification_figures_sub.do		STATA do file
insurance_extensive_margin.do		STATA do file
T4C1.do, T4C11.do, T4C2.do, T4C3.do, T4C4.do, T4C41.do, T4C5.do, T6C2.do, T6C3.do, T6C4.do, T6C5.do, T6C6.do	paramfiles	STATA do file
gmm_age_fv4_pi5_ns_nltax_add	gmm_ado	STATA ado file
gmm_age_fv4_pi5_ns_nltax_b0	gmm_ado	STATA ado file
gmm_age_fv4_pi5_sep_nltax_b0	gmm_ado	STATA ado file
Matlab Program		
Hessian		MATLAB code

- The file `shell.do` can be used to replicate the entire paper. It both defines the path to the program files and data, and calls the rest of the files in the correct order. It is extensively commented explaining what the other files are doing.
- Files `psid_interview.do` and `psid_individual.do` are used to read the PSID interview and individual data and construct the panel.
`psid_consumption.do` prepares household consumption and asset measures.
`psid_sample.do` and `prepare_bootstrap.do` prepare the sample and variables for estimation. The final data set is called `data4estimation.dta`, and it is provided as part of the replication package. For most results there is no need to run these 5 do files again, because `data4estimation.dta` contains the data requires for estimation.

Note 1: The PSID files are not provided. Researchers interested in constructing the data from scratch can download the PSID files (at no cost) from the PSID website and run the do files from the point above on the PSID annual files.

Note 2: Some of the results are generated using extended samples that are not included in `data4estimation.dta`. These results include Columns 3-6 of Table 1 (descriptive statistics), Table 2 (Comparison of PSID data with NIPA) and Column 2 of Table 8. Researchers interested in replicating these results will use the data set `psid_cons.dta` provided in the output subdirectory. They will need to run on this data set the `psid_sample.do` followed by the `prepare_bootstrap.do` files, which will generate the required intermediate data set.

- `residual_measures.do` generates the residual measures for the estimation.
- `advanced_information_test.do` conducts the advanced information test.
- `bootstrap.do` is the file used to conduct the structural estimation. This file will call `variance_estimation.do` (estimates variances), `tax_rate_estimation_snap.do` (estimates the tax function), `pi_s_estimation.do` (projects π and s), and `frisch_estimation.do` (estimates the Frisch elasticities). Note that `frisch_estimation.do` will call the correct ado file from the `gmm_ado` directory. Running the configuration T4C11 will also generate Figure 1 in the paper.

Note 3: To replicate the entire paper the `bootstrap.do` file should be ran 12 times for 12 different specifications (covering all the specifications in Tables 4 and 6, plus two runs for Table 3). This is explained in detail in the comments at the header of `bootstrap.do`.

- `aggregate_consumption.do` and `descriptive_stats.do` generate Table 2 and Table 1 respectively. Note that both of these files require data not in `data4estimation.dta` (follow instructions in Note 2 to run these files).

- `tables_7_8_9.do` generates Table 7, 8 and 9 from the paper. Column 1 of Table 8 requires data not in `data4estimation.dta` (follow instructions in Note 2 to run these files).
- `pi_s_by_age_figure.do`, `marshallian_figure.do`, and `identification_figures.do` generate Figures 3-7 in the paper, as well as the Figure in the Appendix. `marshallian_figure.do` also provides the required numbers for the calculation of insurance decomposition for the sample of food stamps recipients.
- `calculate_AT_elasticities.do` calculates the after-tax Marshallian elasticities reported in Table 5.
- `moment_fit_table.do` and `external_fit.do` generate the two fit figures (8 and 9) and the appendix fit table.
- `insurance_extensive_margin.do` reports the numbers required for the insurance accounting on the extensive margin.
- `Hessian` calculates the Hessian reports the Hessian and eigenvalues reported in Appendix 5.

Note 4: Other than `Hessian` (only used in Appendix 5), researchers only need STATA to replicate the paper. Note however that the particular options used in the GMM estimation are only available for STATA 13 or higher.

Note 5: The replication package does not include a code for generating Figure 2. This is because this figure is an illustration drawing not based on data.

All the program files contain enough explanatory notes so that replication should be easy, although replicating all bootstrap specifications in Tables 4 and 6 might take a few days (depends on computing power).

The files have been checked for bugs, but of course it is possible that we may have missed some (hopefully not!). Please contact us at r.blundell@ucl.ac.uk, pista@stanford.edu or itaysap@post.tau.ac.il if you have any questions or comments.