This folder contains replication files for "Dynastic Precautionary Savings".

Data

The datasets used in the analysis are:

1. PSID (Panel Study of Income Dynamics)

   The PSID data is publicly available at `https://simba.isr.umich.edu/data/data.aspx`. PSID does not allow the transfer of public data that has been downloaded from the website to any third parties.To download the data, one has to first register by completing a registration form and choose a username and password that allows access to the public use data archive. The path to download the data is: Home → Data → Packaged data → Main and Supplemental Studies. The paper uses the Family files, the Cross-year individual file, the Childbirth & adoption history file, the Parent identification file, the 2013 Family rosters and transfers file, and the Wealth supplements.[1]

2. Consumer Expenditure Survey

   The CEX data is also publicly available, but I use the dataset compiled by Mark Aguiar and Erik Hurst and used in Aguiar and Hurst (2013). This is available as supplementary material at `https://www.journals.uchicago.edu/doi/suppl/10.1086/670740`. The dataset is also in the folder Data/CEX.

Code

The construction of the final dataset uses Stata, Fortran and Matlab and proceeds in the following steps:

1. **Construct taxes paid using TAXSIM**
   The programs are in the folder TaxSim. The master file is taxsimPSID_Master.do and it calls all the other programs in the folder.

2. **Construct PSID panel**
   The programs are in the folder Panel. The master file is Master.do and it uses all the other programs in the folder in the following order. Some of the Stata files use the commands winsor and unique which can be installed using the command ssc install winsor and ssc install unique. It first imputes consumption in PSID based on the demand equations estimated with CEX data. The imputation procedure begins with the CEX sample selection performed by CEX_sample.do. The PSID sample selection and some corrections are done in PSID_sample_fam.do, PSID_sample_ind.do and PSID_sample_adj.do. Finally, the imputation is done in impute.do. The PSID variable selection and injection of consumption and tax information is done in PSID_select_variables.do. The panel is constructed in PSID_panel.do and cleaned further in PSID_panel_clean.do.

3. **Consumption age profile**
   The programs are in the folder Facts. The master file is Master.do and it calls the

[1]Note: The variables in the wealth supplements are now also part of the Family files.

program PSID_consumption.do, which estimates the age profile of consumption. The results are then read by the Matlab program age_profiles.m, which generates Figures 3, 4 and 9. The figures are saved as .eps files in the folder Facts/Figures.

4. **Estimate income uncertainty**
   The programs are in the folder Panel. The master file is Master.do and it uses all the other programs in the folder in the following order. The program PSID_uncertainty.do estimates the earnings projections and exports the residuals in the folder Panel/Outsheet. The program PSID_pi.do estimates earnings projections and exports predicted earnings.

   The permanent income uncertainty in Figures 1, 2, 5-8 is constructed using the programs in the folder Uncertainty. The sub-folder Fortran Code contains the Fortran programs to calculate permanent income uncertainty. The main program is _Estimate.f90 and it calls all other Fortran programs in the folder. The Fortran code can be compiled using the file Makefile or can be set up as a project in Microsoft Visual Studio. Running the Fortran code requires the IMSL library. The input used is that produced by PSID_uncertainty.do above and has to be moved manually from the folder Panel/Outsheet to the folder Uncertainty/Fortran Code/Input. The Matlab program Uncertainty_graphs.m reads the output of the Fortran program and generates Figures 1, 2, 5-8, as well as the files with estimates of permanent income uncertainty that are merged into the PSID as described below. This Matlab file uses the Statistics and Machine Learning Toolbox. The figures are saved as .eps files in the folder Uncertainty/Figures.

   The expected permanent income used as control in the regression analysis is constructed using the programs in the folder PI_long. The sub-folder Fortran Code contains the Fortran programs to calculate permanent income. The main program is _Estimate.f90 and it calls all other Fortran programs in the folder. The Fortran code can be compiled using the file Makefile or can be set up as a project in Microsoft Visual Studio. Running the Fortran code requires the IMSL library. The input used is that produced by PSID_pi.do above and has to be moved manually from the folder Panel/Outsheet to the folder PI_long/Fortran Code/Input.

   Standard errors in the paper are bootstrapped. To that end, the program PSID_boot.do in the folder Panel repeats the procedure in PSID_uncertainty.do and PSID_pi.do for all bootstrapped samples. The results would then have to be used as input in the Fortran programs described above, which then have to be ran as many times as bootstrap samples. Note that this process is not automatized. The Matlab program Uncertainty_graphs_boot.m writes files with estimates of permanent income uncertainty based on bootstrapped samples that are then merged into the PSID as described below.

5. **Construct parent-child pairs**
   The programs are in the folder Pairs. The master file is Master.do and it uses all the other programs in the folder in the following order. It first establishes the parent-child pairs in PSID_pairs.do. It then injects income uncertainty information in the sample of PSID parents in PSID_parents.do and PSID non-parents in PSID_non_parents.do.

6. **Estimate regressions**
   The programs are in the folder Pairs. The master file is Master.do and it uses all the other programs in the folder. The program PSID_analysis.do estimates the specifications

summarized in Tables 1-4, 8 and 9. The program generates the point estimates and robust standard errors. The file results.log shows the output of this program. The results in Table 1 are generated by lines 201-203, 267-278. The results in Table 2 and Table 8 are generated by lines 346-373. The results in Table 3 and Table 9 are generated by lines 406-422. The results in Table 4 are generated by lines 378-393. The program PSID_analysis_boot.do injects into the parent-child pairs sample income uncertainty information calculated based on the bootstrapped samples and re-estimates the specifications that form the basis of Tables 1-4, 8 and 9 as many times as bootstrapped samples. Finally, the program PSID_analysis_boot_summary.do summarizes these results.

References

Aguiar, Mark and Erik Hurst. 2013. "Deconstructing Life Cycle Expenditure." *Journal of Political Economy* 121(3):437–492.

PSID. Panel Study of Income Dynamics. public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.