## Overview of files and folders in .\basics_rel_demand

Files in this folder are used to create summary tables and figures, as well as all analysis related to relative demand, in the paper and appendix

All folder references below begin with root folder: .\FINAL_Replication_CLMP_JPE

### Code

- *main.do*
  - sets folder structure and executes all other .do files in this folder.
  - Must set $root folder.

- *1_fixed_effect_fin.do*
  - reads in panel of mothers and fathers and estimates fixed effects for log wages

    The loop over mothers and fathers:

    A. Reads in the panel of mothers (and fathers for married mothers): .\data\data_derived\MotherPanelCDS.dta
    B. Introduces sample restrictions.
    C. Estimates fixed effects (FE) for mother's and father's log wages
    D. Saves files with FE that will be matched to the child panel by mothers and fathers' ID: .\code\estimation\basics_rel_demand\temp_data\ln_HCm_dat.dta and .\code\estimation\basics_rel_demand\temp_data\ln_HCf_dat.dta

- *2_create_child_panel_fin.do*
  - merges the panel of CDS children with additional datasets: predicted wages from census, prices, fixed effects

    A. Reads-in the panel of children: .\data\data_derived\ChildPanelCDS.dta
    B. Renames variables and generates some age-related variables
    C. Merges predicted wages for fathers and mothers from Census (.\data\data_derived\predicted_lnwage3.dta) with child panel
    D. Merges input prices (.\data\data_derived\p_extendedcosts.dta) with the child panel and deflates them to 2002 dollars
    E. Merges parents' log wage fixed effects created above (ln_HCm_dat.dta and ln_HCf_dat.dta) with the child panel
    F. Saves dataset: .\code\estimation\basics_rel_demand\temp_data\psid_child.dta

- *3_data_cleaning_child_panel_fin.do*
  - data cleaning and variable creation to construct the final estimation dataset

    A. Reads in psid_child.dta created above
    B. Performs sample selection
    C. Main data cleaning and construction for childcare expenditures, time investment, goods expenditure, log relative input price and expenditure ratios, indicators for living in the state/region of birth for household head
    D. Drops observations that do not satisfy sample restrictions and saves data for production function estimation and counterfactual analysis: .\data\data_derived\psid_fam.dta and .\data\data_derived\psid_fam.csv

- *4_summary_fin.do*
  - Produces figures with summary statistics on child investment inputs

    A. Reads in data for kids with reasonably aged parents from psid_fam.dta created above
    B. Generates expenditure shares for time, goods, and childcare investment
    C. Introduces final sample restrictions
    D. Constructs the following summary graphs

    - *Appendix Figure G-1: Weekly child investment expenditures by mother's education, includes families with zero child care spending (PSID 2002)*
    - *Appendix Figure G-2: Expenditure shares by mother's education, includes families with zero child care spending (PSID, 2002)*
    - *Figure 1: Weekly child investment expenditures by child's age (PSID, 2002)*
    - *Appendix Figure G-4: Weekly child investment expenditures by number of children and mother's education (PSID, 2002)*
    - *Figure 2: Weekly child investment expenditures by mother's education (PSID, 2002)*
    - *Figure 4: Expenditure shares by mother's education (PSID, 2002)*
    - *Appendix Figure G-3: Expenditure shares by child's age (PSID, 2002)*
    - *Appendix Figure G-5: Expenditure shares by number of children and mother's education (PSID, 2002)*
    - *Figure 3: Weekly investment hours by mother's education (PSID, 2002)*

- *5_demand_estimation_fin.do*
  - performs main relative demand estimation

    A. Reads in data for kids with reasonably aged parents from psid_fam.dta created above
    B. Introduces sample restrictions
    C. Introduces local lists of regressors and variables used in the estimation
    D. Estimates predicted probabilities of work for parents; performs preliminary analysis and computes summary statistics; computes residual variances for Monte Carlo analysis; saves the final estimation sample for relative demand estimation in .\code\estimation\basics_rel_demand\temp_data\allfam_tmp.dta; produces the following tables:

- *Appendix Table G-2: Predicted probability (average derivatives) of work probits for parents*
- *Appendix Table G-3: Log wage regressions for parents*
- *Appendix Table G-13: Probit estimates for positive child care expenditures*
- *Table 1: Summary statistics for full sample: 2002 and 2007*
- *Appendix Table G-1: Summary statistics for restricted samples: 2002 and 2007*
- *Appendix Figure G-6: Distributions of log relative input prices (PSID, 2002 and 2007)*

E. Estimation of time vs. household goods relative demand (2002, 2007); produces the following tables:

- *Table 2: OLS and 2SLS estimates for mother time/goods relative demand*
- *Appendix Table G-5: 2SLS estimates for mother time/goods relative demand using different predicted wage measures as instruments*
- *Appendix Table G-4: First-stage estimates for mother time/goods relative demand using different predicted wage measures as instruments*
- *Appendix Table G-6: OLS & 2SLS estimates for mother time/goods relative demand for different selection on predicted probability of work*
- *Appendix Table G-10: OLS estimates for mother time/goods relative demand by child age*
- *Appendix Table G-12: OLS estimates for mother time/goods relative demand conditioning on 1997 AP and LW scores*
- *Appendix Table G-11: OLS estimates for mother time/goods relative demand by father's wage*
- *Appendix Table G-9: Heckman two-step estimates for mother time/goods relative demand (two-parent households)*
- *Table 3: OLS estimates for parental time vs. goods relative demand, by parent type*
- *Appendix Table G-7: 2SLS estimates for parental time vs. goods relative demand*
- *Appendix Table G-8: OLS estimates for parental time vs. goods relative demand including parental log wage fixed effects, by parent type*

F. Estimation of childcare vs. household goods relative demand (2002, 2007); produces the following tables:

- *Appendix Table G-15: OLS estimates for child care/goods relative demand*
- *Table 4: OLS estimates for child care/goods relative demand, all families*
- *Table 4: columns 1-4*
- *Table E-1: column 5*
- *Appendix Table G-14: Heckman two-step estimates for childcare/goods relative demand*
- *Appendix Table G-16: OLS estimates for child care/goods relative demand by child age*

**Other folders**

- .\code\estimation\basics_rel_demand\logs – Contains log files produced by .do files described above

- .\code\estimation\basics_rel_demand\temp_data – Used as a repository for temporary data files when running Stata .do programs

**Important data sources**

- ./data/data_derived/p_extendedcosts.dta
    - state-year panel of childcare and goods/services input prices

- ./data/data_derived/ChildPanelCDS.dta
    - panel of children in PSID/CDS

- ./data/data_derived/MotherPanelCDS.dta
    - panel of mothers in PSID

- ./data/data_derived/predicted_lnwage3.dta
    - predicted wages based on census data

**Data created for GMM estimation and counterfactual analysis**

- ./data/data_derived/psid_fam.dta/csv
    - main PSID sample with cleaned variables and basic sample restrictions