Overview of Data Files and Programs for "Intergenerational Effects of the EITC: The Case of Grandparents"

Sara LaLumia

1 Datasets for Analysis

All of the analysis in the paper can be created using these cleaned datasets:

- OrganizedOldGenData
- OrganizedParentLevelData
- OrganizedOldGenData_CPS
- OrganizedEmptyNesterOldGenData

2 Creating Tables and Figures

- NTJ_MakeTables.do creates Tables 1 through 7.
- NTJ_MakeFigure1.do creates Figure 1.
- NTJ_MakeFigure2.do creates Figure 2.
- NTJ_MakeFigure3.do creates Figure 3.

3 Assembling Datasets

The raw SIPP files can be downloaded from the NBER website. The raw SIPP files are used as inputs into the following do files. These do files should be run in the order listed, because outputs created in some of the earlier do files are used as inputs in later do files.

1. The do files **NTJ_Assemble_FamilyVariables_SIPP.do** and **NTJ_Assemble_FamilyVariables_SIPP_2014.do** assemble information on household structure. Create variables indicating whether a household includes any potential EITC-eligible children (children under age 19, or children who are full-time students and under age 24), whether a household contains any parents or any mothers of those kids, and whether a household includes any parents of those parents. Create data extracts including relevant information on everyone who meets the definition of a middle-generation parent in a three-generation household, everyone who meets the definition of an oldest-generation grandparent in a three-generation household, and everyone who is an "empty nester," a parent whose own children aren't living in the household and who is living with an own parent.

   Outputs of these files are

- MidGenParentDataYYYY_wx where $YYYY$ indicates the year of the SIPP panel and x represents the wave of that panel. These files include one observation per month per middle-generation parent living in a three-generation household.
- OldGenParentDataYYYY_wx files include one observation per month per oldest-generation grandparent living in a three-generation household.
- EmptyNesterDataYYYY_wx files include one observation per month per middle-generation parent living in a household with one of her parents, but not currently living with any of her own children. This group is used to identify those in the falsification exercise of Table 7.
- EmptyNesterDataGPYYYY_wx files include one observation per month per oldest-generation grandparent living in a household with one of her own children, but not currently living with any of her own grandchildren. This group is used in the falsification exercise of Table 7.

2. The do file **NTJ_Assemble_TopicalModuleWellBeing.do** extracts variables from topical modules.

   The outputs of this do file are

   - ChildCareVars_YYYY_wx files include information on grandparent-provided child-care.
   - WellBeingVarsYYYY_wx files include either health status variables, financial variables, or both (depending on the wave).
   - FertilityHistoryVarsYYYY_wx files include information from the fertility history modules on how many children a person has ever had.

3. The do file **NTJ_Assemble_Kid_Counts.do** puts together information on number of own kids and age of youngest child for each mother and father.

   The outputs of this do file are

   - KidCountYYYY_wx files include person-month information about number of kids of any age, and number of kids in certain age brackets.

4. The do file **NTJ_OrganizeParentLevelData.do** organizes data in a format that is ready for analysis of middle-generation parents. It includes a calculation of the max EITC amount each middle-generation parent could be receive, based on that person's number of kids and state of residence. It also calculates the sum of max EITC amounts that could be received by all middle-generation mothers within a household.

   The outputs of this do file are

   - OrganizedParentLevelData which is ready for analysis of middle-generation parents.
   - MaxMidGenEITCYYYY_wx files contain the maximum EITC amount for any middle-generation parent in the household, ready for merging in with the organized grandparent-level data.

- SumMidGenEITC_YYYY_wx files contain the sum of max EITC amounts for all middle-generation mothers within a household.

5. The do file **NTJ_OrganizeGrandparentLevelData.do** organizes data in a format that is ready for analysis of oldest-generation grandparents.

   The end result of this do file is

   - OrganizedOldGenData

6. The do file **NTJ_OrganizeEmptyNestParentData.do** does some prep work for the falsification exercise of looking at middle-generation parents who are living with their parents, but whose own children don't live in the same household.

   The output of this do file is

   - OrganizedEmptyNesterParentLevelData

7. The do file **NTJ_OrganizeEmptyNestGPData.do** creates the final sample for the falsification exercise.

   The output of this do file is

   - OrganizedEmptyNesterOldGenData which is ready for analysis in the falsification exercise.

8. The following files put together the dataset to carry out analysis of self-reported health in CPS data.

   - NTJ_Assemble_FamilyVariables_CPS.do
   - NTJ_Assemble_Kid_Counts_CPS.do
   - NTJ_OrganizeParentLevelData_CPS.do
   - NTJ_OrganizeGrandparentLevelData_CPS.do

4 Other Files

- **SimulatedEITC** includes a simulated federal and state EITC amount for cells defined on the basis of year, state, marital status (married or head of household), and number of dependent children (one, two, or three plus).
- **OrganizedStateControls** includes annual information for each state on the unemployment rate, the maximum monthly AFDC/TANF benefit amount for a 3-person family, the top state income tax rate, the minimum wage, the gross state product, and an indicator for whether any welfare waivers were in place.