**Panel Study of Income Dynamics (PSID)**

PSID is a nationally representative household panel survey started from 1968. PSID families are followed over time, and households formed by their children and grand children are also invited to join the survey, which allows us to construct genealogical linkages between survey respondents even if they are not living in the same household. More information about PSID can be found: <https://psidonline.isr.umich.edu/default.aspx>.

**Description of the family dataset**

The sample population of this dataset is all survey respondents in the 2013 PSID, a total of 24,759 individuals from 9,040 households. For each individual respondent, survey ID and age of their parents, children, and siblings are included if available. The 1968 ID uniquely identifies survey individuals across years and a zero value means missing. Parents, children, or siblings who were known to have deceased in or before 2013 were coded missing in their 1968 IDs. The 2013 family ID identifies current household unit in the 2013 PSID. Age variables were consolidated from various data files.

The weighted mean age of the sample population is 37.47. On average, each individual of the sample has 1.32 child and 1.37 sibling.

**Data dictionary**

|  |  |  |
| --- | --- | --- |
| Name | Description | Note |
| v1 | Individual ID | 1968 ID |
| sstrata | Sampling error stratum |  |
| scluster | Sampling error cluster |  |
| pwtlong | 2013 individual longitudinal weight |  |
| pwtcross | 2013 individual cross-sectional weight |  |
| fuid | 2013 household ID | First four digits of 2013 individual ID |
| sex | 1- male; 2- female |  |
| agei | Individual age in 2013 |  |
| vm | Mother ID | 1968 ID |
| agem | Mother age in 2013 | Value of zero means deceased in or before 2013; ID coded as zero |
| vf | Father ID | 1968 ID |
| agef | Father age in 2013 |  |
| num\_chd | Number of children |  |
| vc1~19 | Child ID | 1968 ID |
| age\_chd1~19 | Child age in 2013 |  |
| sex\_chd1~19 | Sex of child |  |
| num\_sib | Number of siblings |  |
| vs1~13 | Sibling ID | 1968 ID |
| age\_sib1~13 | Age of sibling |  |
| sex\_sib1~13 | Sex of sibling |  |
| sibtype1~13 | Type of sibling:  1 "full sibs" -- biological mom and dad are known for both  sibs, and both parents are in common  2 "half sibs, mom in common" -- biological mom and dad are  known for both sibs, and only the mom is in common  3 "half sibs, dad in common" -- biological mom and dad are  known for both sibs, and only the dad is in common  4 "at least half sibs, mom in common" -- biological mom known  for both sibs (but bio dad not known for both sibs), and the  sibs have the same bio mom  5 "at least half sibs, dad in common" -- biological dad known  for both sibs (but bio mom not known for both sibs), and the  sibs have the same bio dad  6 "all others" (excluded) |  |

**Creation of the dataset from 2013 PSID**

All survey respondents in 2013 PSID were selected from the cross-year individual file to serve as the base dataset.

To construct family relationships, the base dataset was merged with the parent identification data (PID) that links individuals to their birth parents and children. The PID file consolidates information from birth history data as well as various other sources to crosswalk children and their parents’ identifiers. Since not all data sources differentiate between biological and adoptive parents, the PID file assumed all child-parents relationships to be biological unless contradicted by any differentiated sources. The Family Identification Mapping System (FIMS) data was used to obtain pairs of siblings and their identifiers. Only siblings who at least share the same biological mother or father were included.

Obtaining demographics of parents or children or siblings has a major data limitation. It is straightforward to use the cross-year individual file if the person has ever been a survey respondent. However, the main survey does not necessarily follow everyone in the family. For example, a new individual can join the survey because of marriage to a PSID individual but that person’s parents and children from previous marriages are not followed or traced back. In such case, the person may have never been included in the main survey. In a less direct way, the parent birth year variables in the cross-year individual file (from the Childbirth and Adoption History File) and the parent age variables in the 2013 Family Rosters and Transfers File were used to partially recover missing data. The Family Rosters and Transfers File is particularly relevant because it asked about parents and children of household heads (and spouses) in 2013 PSID regardless of their eligibility for the main survey. Survey questions in 2013 were the most preferred sources to obtain age information. If not available, ages were deducted from data in other years if possible. People known to have deceased in or before 2013 according to the cross-year individual file were excluded.