PSID SAMPLE: AMERICA'S FAMILY TREE ⧉

The original PSID sample of roughly 18,000 people in 5,000 households consisted of a nationally representative sample and an oversample of low-income families. The oversample was included to facilitate investigations of poverty-related issues. PSID data include economic, social, and health information collected using in-person, telephone, computer-assisted interviewing methods, and a web option. PSID families are followed regardless of where they live. The sample grows naturally as children and grandchildren from these families form their own households and are invited to join the PSID. Samples of immigrants have been added to the PSID in 1997/1999 and 2017/2019 so that the sample continues to closely resemble the national population, making PSID America's Family Tree.

CONTENT OF CURRENT QUESTIONNAIRE ⧉

Employment	Education
Wages & Income	Marriage & Fertility
Expenditures	Health Status
Wealth	Health Behaviors
Mortgage Distress & Foreclosures	Health Insurance
Pensions	Program Participation
Philanthropy	Computer Use & Technology
Time Use	Housing Characteristics

STUDY DIRECTION AND OVERSIGHT ⧉

PSID is directed by faculty at the University of Michigan, with data collection conducted by the Survey Research Operations group within the Survey Research Center at the U-M Institute for Social Research. An external Board of Overseers, comprised of leading scientists from various disciplines, provides scientific input to the project.

SPONSORSHIP ⧉

Major sponsorship of the PSID is provided by: the National Science Foundation, the National Institute on Aging, the Eunice Kennedy Shriver National Institute on Child Health and Human Development, the Center on Philanthropy at Indiana University, and the United States Department of Agriculture.

DATA ACCESS: PSID.org ⧉

All waves of the PSID and study supplements are accessible via the PSID Data Center, which provides:

- Automatic merges of all waves of data from the PSID and its major supplements
- Customized codebooks and data extracts in a variety of formats including SAS, Stata, SPSS, Excel, and text
- Instructional web tutorials, bibliographies, technical papers, and user documentation

Some types of PSID data are available only under a restricted contract. These include but are not limited to: geospatial data below the level of state; mortality data; assisted housing data from the U.S. Department of Housing and Urban Development; Medicare claims; and educational characteristics from the National Center for Education Statistics.

RESTRICTED USE DATA

PANEL STUDY OF INCOME DYNAMICS

PSID

PSID.org ⧉

PSID OVERVIEW ⧉

The PSID is the world's longest running nationally representative household panel survey. With over 50 years of data on the same families and their descendants, the PSID is a cornerstone of the data infrastructure for empirically based social science research in the U.S. The PSID gathers data on the family as a whole and on individuals residing within the family, emphasizing the dynamic and interactive aspects of family economics, demography, and health. PSID data were collected annually from 1968-1997 and biennially after 1997.

With low attrition and high success in following young adults as they form their own families, the sample size has grown from roughly 5,000 families in 1968 to more than 9,200 families and 24,000 individuals by 2021. The PSID has distributed data on more than 84,000 individuals over the course of the study, collected through 2023.

RESTRICTED DATA OVERVIEW ⧉

To safeguard the confidentiality of the respondents at the highest level, some data are provided only under conditions of a restricted contract. A variety of restricted data are available including: geospatial data; mortality data; assisted housing data from the U.S. Department of Housing and Urban Development; Medicare claims; and educational characteristics from the National Center for Education Statistics. Contractual arrangements require review of a research statement, acceptable data use policy, IRB approval, and an administrative fee. Access to PSID restricted data may be obtained via secure remote access through a Virtual Data Enclave. For more information, visit our website at psid.org.

GEOCODE DATA ⧉

The geocode data files include the ID codes required to link data from the PSID in each wave to contextual data from secondary sources such as the Census. Geospatial data below state level including County, ZIP code, Metropolitan Statistical Area (MSA), Core Base Statistical Area (CBSA), Census place, tract, block group, and block are available.

MORTALITY DATA ⧉

The Mortality File contains information about the deaths of all individuals in the PSID who are known to have died from the beginning of the study through the most recent wave of data collection (n~7,800). Variables include the year in which PSID staff uncovered death information, month/year of death, age at death, location of birth and death, and an accuracy indicator for date of death.
Additional variables from the National Death Index (NDI) are included, such as complete cause of death information starting in 1979 for those cases deemed to be a good match. The data also include month and year of death as well as identifiers that allow users to link the mortality data to the publicly available data

SCHOOL IDENTIFIER DATA ⧉

School identifiers from the National Center for Education Statistics (NCES) Common Core of Data (CCD), Private School Survey (PSS), and Integrated Postsecondary Education Data System (IPEDS) have been linked to individuals in the PSID, the Child Development Supplement (CDS), and Transition into Adulthood Supplement (TAS).

The CCD is a national database of all public elementary and secondary schools and school districts in the U.S. and contains general descriptive information on schools and school districts, students and staff, and fiscal data. The PSS is the counterpart for all private schools in the country. IPEDS collects data at the post-secondary school level on enrollments, graduation rates, faculty, staff, finances, costs, and student aid.

ASSISTED HOUSING MATCH DATA ⧉

Each PSID family in every year has been identified as living in housing units subsidized by HUD, the Farmer's Home Administration, through tax credits administered by the Department of Treasury, or state housing programs. This was accomplished by address matching of PSID families with those in the Assisted Housing Database (AHD). The data provide an indicator for whether a PSID address available for a given year corresponds to an assisted housing address, and, if so, the type of assisted housing (e.g., public housing, Low Income Housing Tax Credit, state subsidy).

PSID-CRCS DATA ⧉

The Childhood Retrospective Circumstances Study (PSID-CRCS) is the first study conducted by the PSID using the internet as the primary mode of data collection. The goal of the study was to design and collect a mixed mode (web or paper) module from reference persons and, if married/cohabiting, spouses/partners, about their childhood experiences.

GENOMIC DATA ⧉

CDS-2014 collected saliva samples from children and their primary care givers. These samples were used to obtain three components of genomic data, including publicly available polygenic scores, and restricted access to a genetic relatedness file and a Variant Call Format file. which includes the full set of genetic sequence data.

MEDICARE AGGREGATED CLAIMS DATA ⧉

Linkages to over 20 years of Medicare claims have been made available. Claims data begin with a respondent's first service after Medicare entitlement, which is at age 65 or when qualification for Disability Insurance is met. Claims files include short-stay inpatient, skilled nursing facilities, hospital outpatient, physician/supplier, home health, and hospice services. Summary files are compiled annually and by quarter. Individuals who wish to obtain these data must obtain Data Use Agreement (DUA) from the Centers for Medicare & Medicaid Services (CMS) as well as approval from PSID.