



HHS Public Access

Author manuscript

Ann Am Acad Pol Soc Sci. Author manuscript; available in PMC 2019 October 30.

Published in final edited form as:

Ann Am Acad Pol Soc Sci. 2018 November ; 680(1): 9–28. doi:10.1177/0002716218809363.

Fifty Years of the Panel Study of Income Dynamics: Past, Present, and Future

David Johnson, Katherine McGonagle, Vicki Freedman, Narayan Sastry

University of Michigan

Abstract

The Panel Study of Income Dynamics (PSID) is the world's longest running household panel survey. Beginning in 1968, it has collected data on the same families and their descendants, making it an essential part of America's data infrastructure for empirically based social science research. PSID arose from the War on Poverty as a tool for evaluating poverty dynamics, and this year marks 50 years of data collection. Because of its long history and distinctive design of following adult children as they form their own households, PSID is uniquely positioned to address emerging social and behavioral research questions and related policy issues. This overview presents the design and structural aspects and its evolution over the past 50 years, the successes of the current survey, possible future directions, and the value of using the PSID to understand the challenges facing American families.

Keywords

panel study; life course; data collection; intergenerational

“Won't you be my neighbor?” was the often heard line from the title character, Mr. Rogers, in the popular children's television show “Mister Roger's Neighborhood” that debuted in 1968. That same year, the Panel Study of Income Dynamics (PSID) began its remarkable history of collecting data about families, neighborhoods and people, bringing it all together in one place. With 40 waves of data collected on the same families and their descendants as of 2017, PSID is a cornerstone of the data infrastructure for empirically based social science research in the United States. Its long-term measures of economic and social well-being of a representative sample of U.S. families have allowed researchers and policy makers to study the dynamics inherent in social and behavioral processes. 50 years on, PSID is now the world's longest-running household panel survey and has become, for all intents and purposes, a data portrait of America's Family Tree.

PSID was originally created to assess the impact of President Lyndon Johnson's War on Poverty. In 1966 and 1967, the Office of Economic Opportunity (OEO) directed the U.S. Bureau of the Census to design and field the Survey of Economic Opportunity (SEO) to provide data for a national assessment of War on Poverty programs. Interest in continuing the study with the primary goal of understanding the dynamics of economic well-being led

OEO to approach the Survey Research Center (SRC) at the University of Michigan about continuing to interview a sub-sample of low-income SEO households. In President Johnson's third annual report to Congress (Office of Economic Opportunity (1968)), he stated the reasons that the SRC was selected to conduct this important study: "SRC has had outstanding success with this type of study...they have now undertaken a study for OEO which will follow a specially selected group of poverty families for several years so that we may gain real insight into the social and economic dynamics of poverty."

This article provides a detailed profile of PSID and its supplemental studies, highlighting features that are valuable to longitudinal and life course research. We also discuss possible future goals for the survey, the important research that has been done using PSID, including that which is shown in the other articles of this volume.

The PSID Main Survey

The primary source of information on PSID sample members is data from a survey conducted annually through 1997 and biennially thereafter. The original aim of studying income and poverty dynamics began with PSID's 1968 sample, which was designed by the SRC: 4,802 households across 40 states, comprising an oversample of 1,872 low-income families from SEO plus a nationally representative sample of 2,930 households.

Professor James N. Morgan, who became the study's first director at Michigan, argued successfully for adding a sample of households from the SRC national sampling frame so that the study would also include non-poor households and hence represent the entire population of the U.S. In addition, a fortuitous decision was made to follow family members who moved out of study households, such as children who came of age during the study. This allowed the sample to remain representative of the nation's families and individuals over time.¹

Approximately 18,000 individuals lived in these original 1968 families and are considered to have the "PSID gene" making them eligible to be followed for subsequent interviews. In addition, all individuals born to or adopted by an individual with the PSID gene acquire the gene themselves, becoming PSID "sample persons" who are followed in the study. As members of sample families grow up, move out, and form their own economically independent households, they are interviewed separately, increasing the overall number of interviews conducted each wave.² This unique design of following children of sample members as they themselves become adults replenishes the sample and helps to maintain its national representation, also facilitating the study of outcomes across generations. In the 2017 wave of the survey, there are some people who are members of the seventh generation of an original PSID family.

¹Detailed information about the study design, following rules, and wave-specific information may be found in the PSID Main Interview User Manual on the PSID website: <https://psidonline.isr.umich.edu/data/Documentation/UserGuide2015.pdf>

²See Table 2 of the 2015 PSID User Manual for the number of original sample members who have been interviewed in each wave and Table 3 which lists the number of sample members in each wave who have left the study, died, or were dropped from the sample: <https://psidonline.isr.umich.edu/data/Documentation/UserGuide2015.pdf>.

In each wave of the survey, PSID respondents provide information about a financially responsible reference person, that person's spouse or partner, and all other family members living together who are referred to as the "family unit." Individuals who are not the reference person or their spouse/partner are classified as "other family unit members" if, at the time of interview, they are members of the respondents' family and either residing in the interviewed family unit or temporarily away in an institution (e.g., college, jail, hospital, or the military). Family members are individuals who are related to the reference person and spouse/partner by blood, marriage, or adoption, and unrelated individuals are also included if they are permanently living together and share income and expenses.

Over 50 years, PSID has collected and made available survey information on 80,666 individuals. The 2017 sample comprised 9,607 families consisting of 14,571 reference persons and spouses/partners and 24,998 other family members. In 2017, 32,990 individuals with the gene had been interviewed for at least 5 waves of PSID. A total of 1,946 people have participated in all 40 waves of PSID, and 1,096 individuals who were under age 18 in 1968 have been interviewed for every wave of the PSID, thus providing information on a substantial portion of their life course (see also McGonagle et al. (2012)).

Over the majority of waves of data collection (i.e., in every wave since 1970, and through 2015), PSID has achieved extraordinarily high wave-to-wave response rates of 94%, and overall response rates of 91%, an accomplishment unparalleled by other panel studies. In recent years, there has been a well-documented rapid decline in response rates for most cross-sectional surveys, and virtually all large, national panel studies are now also experiencing these declines (National Research Council 2013). The reasons are speculative but likely include two major secular changes, both which create a challenging environment for survey research: first is an increase in the time demands of daily life (Bianchi 2011), with a greater proportion of families led by working single parents, and dual employment among coupled individuals; and second, a cultural shift in the manner of communication that has seen greater reliance on the internet and cell phones, and a dwindling number of households (less than 50% of U.S. families as of 2017) with landline phones (Pew Research Center, 2017), and less willingness to answer phones in general.

PSID has been highly responsive to the emerging use of new communication technologies by its sample members. Moving beyond a reliance on traditional landline telephones and postal mail to contact respondents, PSID developed and implemented new protocols based on email (in 2013) and text messages (in 2015). These respondent contact protocols use brief email and text messages to encourage the participation of study members and remind them of upcoming interview appointments, with links embedded in the emails allowing the date and time of the appointment to be placed directly on respondents' calendars.

New technologies such as call-screening and call-blocking have become ubiquitous, substantially increasing the difficulty of contacting respondents to complete their interview. As a consequence, the latest wave of PSID has seen greater fieldwork effort than in prior waves, with an increase in the number of calls by interviewers to make contact with families, higher rates of broken interview appointments, and cell phone use by sample members exceeding 90%. Still, though, when data collection for the 2017 main sample ended on

December 31, 2017, PSID had successfully achieved the study's high reinterview response rate goal of 95%. During data collection in 2017, we expanded the use of text messaging and email to make contact with families, implemented a variety of incentive strategies, and tried new technologies such as an online appointment scheduler and the use of outbound telephone numbers that matched the area-code of respondents. As a result, for the first time, the average number of contacts needed to complete an interview in 2017 dropped slightly.

Sample representativeness.

Despite consistently high response rates, there is evidence that lower income families have higher cumulative attrition (e.g., Fitzgerald et al., 1998; Fitzgerald, 2011). However, parameter estimates of interest have not been found to be biased. In a recent analysis of the effects of cumulative attrition in PSID through 2007, Fitzgerald (2011) finds little-to-no evidence of biased estimates of sibling correlations or of parameters in intergenerational models of health outcomes (see also Schoeni and Wiemers (2015)).

Item nonresponse is also low, with very few questions missing responses for more than 3–4% of cases (Killewald et al., 2011). Sample weights are provided for each wave to account for differential probabilities of selection due to the original sample design and subsequent attrition, including longitudinal individual weights, longitudinal family weights, and cross-sectional individual weights.³

The quality of the data collected in PSID is examined regularly (for examples, see PSID webpage at <https://psidonline.isr.umich.edu/Guide/Quality/DataComparisons.aspx>). There have been over 50 such assessments, covering topics such as income, wealth, poverty, consumption expenditures, health, charitable giving, demographics, sampling and representativeness, event history calendars, and more. A common approach is to compare a particular set of measures in PSID to those in another survey that is regarded as the “gold standard” for that content area, often because that survey devotes more questions to the area than PSID. For example, with the release of each new wave of data, PSID staff conduct appropriately weighted cross-sectional comparisons for household income between PSID and the Annual Social and Economic Supplement of the Current Population Survey. In general, PSID is better at capturing some measures; for example, PSID finds a few percentage points more income at each percentile point than does CPS. Compared to the National Health Interview Survey (NHIS), PSID finds fewer reports of “excellent” current health and somewhat higher rates of work limitations, but very similar rates of specific health conditions, smoking, and obesity (Insolera and Freedman 2017). Cynamon and Fazaari (2015) compare aggregate income measures from CPS, the Survey of Consumer Finances (SCF), and PSID to the National Income and Product Accounts (NIPA). They find that after adjustment for conceptual definitions of “income,” PSID aligns well with NIPA – from 1968 to 2010 the average discrepancy was 3.7% and stable over time, and was characterized as having “the closest correspondence of the three surveys.” Krueger et al. (2015) confirm that both income and consumption measures from PSID track changes in NIPA aggregates.

³Information about the construction of these weights and their analytic use is provided in companion documents on the PSID website (<http://psidonline.isr.umich.edu/Guide/documents.aspx>).

A comparison with the Consumer Expenditure (CE) Survey for expenditures reported in the 1999–2009 waves of PSID (Andreski et al. 2014) concluded that the two surveys find similar levels of overall consumption and similar life cycle patterns. A similar analysis comparing estimates of wealth from PSID with those from SCF reported that PSID found less indebtedness at the bottom of the wealth distribution and less wealth in the top 2–3% in part because SCF expends a great deal of effort to oversample the very wealthy (Pfeffer et al., 2014). Median wealth differs by only 6%.

Content domains.

In addition to collecting information on family income and wealth dynamics since its inception, with input from its Board of Overseers and the broader scientific community PSID has been expanded over time to collect and distribute data in many domains to allow the study topics of emerging scientific and policy interest. Data have been collected on a wide array of social, demographic, economic, and health topics, including employment, earnings, income from all sources, wealth, expenditures, housing, foreclosure and mortgage distress, geospatial data, health status, health behaviors, adult and childhood health conditions, disability, health insurance, marriage and fertility, program participation, vehicle ownership, wealth and pensions, debt, and philanthropy (see PSID User Guide for additional details). Thousands of additional variables that fall into other domains have been collected in various waves throughout the history of PSID. In response to a growing interest in understanding household consumption choices, PSID began expanding the number of questions on consumption expenditures in 1999, with further expansions in 2005.

The most recent enhancements to PSID in the last two waves included enhanced content on internet access, food security, income for spouses/partners to further parallel that of reference persons, health insurance, children, and new sample members' background information and improvements to the wealth and pensions sections, expansion of consumption expenditure questions, addition of stylized time use questions, and the addition of a memory impairment screen. In addition, in response to the growing length of the questionnaire, a continued effort has been made since 2013 to reduce respondent burden while maintaining as much consistency as possible with prior waves. We focused mainly on items that appeared to be challenging (as evidenced in part by relatively long response times) or redundant (as evidenced by modest use of the measures by the research community).

Sponsorship and administration

The National Science Foundation has been the lead funder of PSID over the past several decades. Recent support has also been received from the following organizations: National Institute on Aging, *Eunice Kennedy Shriver* National Institute on Child Health and Human Development, Economic Research Service of the Department of Agriculture, the Center on Philanthropy at Indiana University, the Assistant Secretary for Planning and Evaluation of the Department of Health and Human Services, and the Department of Housing and Urban Development. The study has been housed at the Institute for Social Research, University of Michigan, from its beginning in 1968 to the present and is directed by a team of social

scientists. Since 1982, a Board of Overseers has fostered input from the national community of scholars, researchers, and policymakers.

Supplements and Modules

PSID has conducted supplemental surveys and added new modules to the Core interview in response to emerging scientific issues.

Child Development Supplement.

The Child Development Supplement (CDS) is an ongoing component of PSID funded by the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD) that provides researchers with extensive data on children and their families for studying the dynamic process of early human and social capital formation. The original CDS was launched in 1997 on a cohort of children aged 0–12 in PSID families, and followed these children over three waves, ending in 2007–08 (see McGonagle and Sastry 2015). By 2014, all of the children in this original cohort had reached adulthood, and a new generation of children had replaced them in PSID families. Starting in 2014, a new CDS was launched that shifted the orientation from a cohort study to one that obtains information on the childhood experiences of all children in PSID families at regular intervals of 5 to 6 years. CDS-2014 included all sample children aged 0–17 years and their primary caregivers. The questionnaire content is consistent with earlier waves of CDS to permit cross-cohort analysis of children's development. The resulting data provide researchers with a comprehensive national database of children and their families to study the dynamic process of child development. Topic areas include a broad array of developmental outcomes and factors that influence development such as: physical health, emotional well-being, intellectual and academic achievement, cognitive skills, social relationships with family and peers, time diaries, resources in the home and school, and parenting practices. CDS-2014 also collected saliva samples from children and primary caregivers to obtain genetic indicators, which will create new opportunities to evaluate possible effects of measured genotypes on children's development and relationships between health status and socioeconomic conditions. Preparations for data collection of CDS-2019 are currently underway.

Transition into Adulthood Supplement.

In 2005, PSID began a new supplemental data collection funded by NICHD, called the Transition into Adulthood Supplement (TAS), which followed children from CDS who had turned age 18 and had completed or left high school and had families still active in PSID. TAS bridges the lifecourse period between childhood (as assessed in CDS) and economic independence in adulthood (when sample members become eligible to be interviewed as primary respondents in PSID). TAS has collected data on young adult developmental pathways and outcomes. Topics covered in the TAS interview include: time use (including computer-related activities), psychological functioning, marriage, family, responsibilities, employment and income, education and career goals, health, social environment, religiosity, and outlook on life (see McGonagle and Sastry 2015; Sastry, Fomby, and McGonagle 2018). TAS data on the original CDS cohort have been collected in five biennial waves from 2005 to 2015. Starting in 2017, the TAS underwent a major expansion of the study sample

by following the shift in orientation of the new CDS, collecting information on the experiences of all young adults in PSID families. The TAS instrument was substantially revised, adding new and updated content on topics such as childhood health, fertility-related behavior and pregnancy intentions, and computer skills.

Time Use and Wellbeing.

PSID has expanded its content into time use and wellbeing measurement in several ways. The Disability and Use of Time (DUST) supplement includes not only detailed disability measures but a 24-hour time diary for older adults in PSID, along with measures of experienced well-being and evaluations of satisfaction. When used in combination with PSID's long-term measures of health, economic wellbeing and rich familial context, DUST diary data allow for the examination of factors that promote subjective well-being among older adults experiencing functional loss and those providing assistance. Two waves of DUST have been collected --a sample of older couples in the 2009 PSID and single and married/partnered older adults ages 60 and older in the 2013 PSID. An abbreviated time diary with experienced and evaluative wellbeing measures was also included in the Wellbeing and Daily Life supplement (see below). In 2017, PSID added a stylized time use module to its core interview that asks about time in the last week spent in various activities. Follow-up items focus on physical, cognitive, and social demands of work and leisure activities. The addition of this module to the main survey makes PSID the only national survey with panel measures of a comprehensive set of economic life-cycle budget components: income, consumption, wealth, and time use.

Childhood Retrospective Circumstances Study.

The Childhood Retrospective Circumstances Study (sponsored by NIA) is the first study conducted by PSID using the internet as the primary mode of data collection. Topics included: childhood health conditions, socioeconomic status, neighborhood quality, friendships, school experiences, exposure to the criminal justice system, parent/guardian mental health, relationship quality with parents/guardians, and young adult mentoring. The data are facilitating new studies of the influence of early life conditions on adult health and economic outcomes.

Well-being and Daily Life.

In 2016, the PSID implemented another mixed mode (web or paper) supplement collecting information on wellbeing, personality traits and every day cognitive skills. Content included life satisfaction and flourishing, positive and negative emotions, an abbreviated set of questions to capture time use and wellbeing on the prior day (Lucas et al 2018), personality and self-efficacy, trust-hostility and achievement orientation, verbal reasoning, health literacy, quantitative reasoning, and financial reasoning. Some of the scales (self-efficacy, trust hostility, achievement orientation, verbal reasoning) were also included in the earliest waves of PSID. This new information is enabling expanded research on the interplay of economic, social, and health outcomes with well-being over the life course and within a family context.

Family Rosters and Transfers.

Family structure in the U.S. has changed dramatically since PSID began in 1968. Given these ongoing changes, PSID has a long-standing interest in monitoring whether or not those changes affect the intergenerational transmission of well-being, and the ways in which Americans transfer wealth and resources among family members. This monitoring includes economic resources in later life, such as Social Security, pensions, and other resources. Modeled after PSID 1988 Family Transfers module, with support from NIA, information was collected during the 2013 main PSID interview about each child and parent including age, educational attainment, marital/partner status, and number of children, as well as any recent or long-term transfers of time and money to and from them.

Food Security Module.

A version of the U.S. Household Food Security Survey module developed by the U.S. Department of Agriculture was collected in PSID between 1999 and 2003. To assess whether families remain food secure over time, this module was reinstated in the 2015 PSID and in the 2014 CDS. The questions ask about various levels of food security such as worries about having enough food and enough healthy food, cutting back to conserve food, and running out of money to buy food. These data allow the food security status of each family to be defined along a continuum extending from high food security to very low food security.

Newborn Module.

Starting in 2013, with support from NICHD, a newborn module was added to the main PSID interview that collects information on approximately 800 births per wave. Through 2011, PSID collected information on new births including the child's sex, race/ethnicity, date and place of birth, birthweight, and current place of residence. The new module adds questions about prenatal care, including the number and timing of visits; pregnancy complications and medical and other risk factors for pregnancy, including mothers' tobacco use, pre-pregnancy weight, and weight gain; gestation length; whether the pregnancy was intended; delivery assistance, place and method of delivery, and payment for delivery; characteristics of labor and delivery; health at birth; hospitalization at and after birth; multiple births and whether identical or fraternal; child deceased or adopted out; and ever breastfed.

Enhanced Education Update Series.

A new series of questions about educational attainment and timing for *all* individuals 16+ years of age in PSID families was added with support from NICHD starting in 2013 and collected in each subsequent wave. The new question series support more detailed analyses of the factors that shape educational attainment and timing as well as the effects of schooling on a large number of related behaviors and outcomes. These data also include information about the fields of study, which can be used to identify the characteristics of STEM education and their impacts on a variety of outcomes over the life course.

New Immigrants

A design goal of PSID since its inception has been to represent the national population of U.S. families. In 1997, in response to the rapid growth of the immigrant population since the study began, PSID introduced a refresher sample of post-1968 immigrant families and their adult children. In 2017, PSID received supplemental funding from NICHD that supported the addition of a new immigrant refresher sample that will make PSID sample again representative of the entire U.S. population and will also provide unique opportunities for innovative research on a range of topics. Researchers will be able to study how immigrant families who arrived after 1997 differ from other families, including earlier immigrants (who arrived in the U.S. between 1968 and 1997) and with families who were in the U.S. in 1968 (including immigrants who arrived in the years prior to 1968).

Screening to identify new immigrant families was conducted in collaboration with a major screening effort led by the Health and Retirement Study (HRS) to recruit new cohorts. PSID added two questions to the HRS screener to learn if anyone in the household was foreign born and whether they arrived in the US since 1997. These questions were asked in households in which no one was age-eligible to be part of the new HRS cohorts. The eligible PSID cases were interviewed during the 2017 Core PSID wave, including a new module on family background and language proficiency. Family units were eligible for the new immigrant sample if the primary respondent and/or the spouse/partner (if any) immigrated to the U.S. after 1997 (when PSID last refreshed the sample with new immigrants). At the end of the field period there were 455 completed interviews with new immigrant families at a response rate of 75.5%. All new immigrant families—and their offspring—will be treated the same as existing PSID sample members for future waves.

The collaborative screening effort created a gap for PSID immigrant respondents born between 1960 and 1971 because these respondents were retained by HRS. A multiplicity sampling approach (Sirken 1970) is being used to fill this gap by asking screened-in immigrant families during their 2017 Core interview to identify relatives in the target cohort that have immigrated to the U.S. since 1997.

The PSID new immigrant refresher will also support innovative research on how outcomes among new immigrants vary across countries. Data from PSID and its companion studies in other countries support cross-national research on outcomes such as educational attainment, marriage and cohabitation, living arrangements, and fertility. Several of PSID's companion international studies have nationally-representative samples that incorporate periodic immigrant refresher samples, including the UK Household Longitudinal Study, the German Socio-Economic Panel, and the Household, Income and Labour Dynamics in Australia Survey. Cross-national research has the potential to reveal the effects on child and family outcomes of U.S. immigration policies and circumstances through comparisons with parallel experiences of other countries. For instance, studies can use data from PSID and its companion surveys to examine the pace of immigrant assimilation into society in the U.S. and other countries, and to investigate the effects of assimilation on child and family outcomes.

Data distribution and documentation

Distribution of public use data from PSID – which accounts for the vast majority of all PSID data being distributed – is carried out through a web-based Data Center. The PSID Data Center allows users to create customized longitudinal datasets from all waves of the main interview and supplemental PSID data collections. Users can create customized codebooks specific to the data that have been downloaded, search and browse for variables, and archive data downloads for shared and future use.

PSID creates a variety of generated economic variables in the domains of income, employment, wealth, and consumption. In particular, income and employment components for primary respondents and their spouse/partners include taxable, transfer, and social security income, earnings and weeks worked (see Duffy 2011), and wealth generated variables include business equity, transaction accounts, debt, equity in real estate and primary residence, stock, vehicles, IRAs, and other assets (see PSID User Guide 2017).

In addition, because of its genealogical design, PSID is one of the few nationally representative U.S. datasets that can be used for intra-and intergenerational analyses. To help users with these important but complex relationships, PSID offers the “Family Identification Mapping System” (FIMS) software tool, which easily identifies relationships within and across generations. FIMS creates a customized file – i.e., “map file” -that contains the identification variables of the relatives an analyst wishes to examine. For example, an analyst could choose “biological grandparents” as the relative of interest. FIMS would then generate a file that contains the IDs of all PSID sample members and each of their biological grandparents present in PSID study. A comprehensive User Guide for FIMS is available on PSID website.

The PSID Data Center includes a comprehensive cross-year variable index that allows users to browse, search, and download all public data in the archive. The index provides users with a view of a given domain of variables – income, health, or wealth, for example – which they can then “unfold” to view all of the variables related to that domain. For a given specific variable – for example, current employment status – the index will list the years that the variable is available. Users can then click on the year to view the codebook for that specific year, and then click on the year indicator for the given variable to add the variable for that year to a data cart.

To maintain respondent confidentiality, some PSID data are provided only under the conditions of a restricted data use agreement between the researcher and the University of Michigan. In order to access these data, users must enter into a data use contract and agree to an appropriate data safeguarding and disclosure plan. The restricted data are available through remote access to a secure virtual data enclave at the University of Michigan. These restricted data include geospatial identifiers, a subset of sensitive data, and a variety of linked administrative records. Geospatial identifiers, including Census tract, block-group, and block, are available for the current residence at each wave for all families since 1968. County-level data on where PSID individuals and their parents grew up are also available. Sensitive data include information on sexual activity, contraception, pregnancy and births,

delinquent behavior, and tobacco, drug, and alcohol use (collected from CDS children 12 years of age and older); exposure to the criminal justice system (collected in the TAS and CRCS); and childhood health and parental mental health and relationship quality (collected in the CRCS). Information on decedents is collected in each wave of the main PSID interview, including dates and location of deaths. Links to external administrative data sources have been expanded in recent years, with information now available on: school characteristics of primary, secondary, and postsecondary public and private schools from the U.S. Department of Education's National Center for Education Statistics (NCES) Common Core of Data (CCD), Private School Universe Survey (PSS), and Integrated Post-secondary Education Data System (IPEDS); assisted housing identifiers from the U.S. Department of Housing and Urban Development; and health care claims spanning 1991–2010 from the Centers for Medicare and Medicaid Services.

Research potential of the PSID

The extended time series and consistently high response rates of PSID make it a substantial and powerful analytic tool for studying the antecedents and consequences of a range of social, health, and economic conditions in the U.S. Since individuals are followed over the entire life course, the number of waves that an individual appears in the sample is related to their age. However, individuals who were born into PSID families also have substantial information on their childhood circumstances from their parents' and grandparents' reports. This combination of information on childhood circumstances and later adult behavior and outcomes represents a major analytical strength that supports a variety of different analyses. Moreover, the analytical samples are large. For example, 2,209 individuals who were under age 18 in 1968 were age 50 or older by 2017, and are therefore represented in the data for almost 50 years of their lives. This cohort, born 1950–1968, overlaps substantially with the 1946–1964 Baby Boom generation and is now reaching the ages where life events such as acute health shocks, chronic disease, and labor force transitions occur. For children in the CDS and TAS, a wealth of information exists on their behavior, health, and development as they passed through early, middle, and late childhood and into adolescence and young adulthood, and, increasingly, as members of the main PSID. The resulting data can be used to better understand how early circumstances, such as childhood socioeconomic status, health, neighborhood and school characteristics, and educational choices shape health and well-being over the life course.

The following of sample members who are descendants of the original 1968 sample yields intergenerational and intragenerational samples which are a powerful addition to the study of life course development. The intergenerational sample is comprised of adult children who split off from their sample parents to form their own households and who are recruited into the study. Relative-pairs that can be examined in PSID include dyads formed from parents and children, from siblings, and from grandparents and grandchildren. There are hundreds, and in many cases thousands, of these pairs within PSID; they are typically observed for many years, leading to a large number of relative-pair-year observations. In the 2015 wave, there were 4,822 parent-child family pairs and 1,371 grandparent-grandchild family pairs in the sample, where both individuals are financially responsible for their own family unit. For

many PSID families, self-reported information is currently available on up to five generations within the same family at various points in their life course.

PSID has been one of the most widely used social science data sets in the world, and the rate of use has been increasing in recent years. One indication of the importance of PSID was the naming of the study to the National Science Foundation's "Nifty 50", which were the 50 most significant NSF-funded studies in the first 50 years of NSF's existence. PSID was the only social science study on the list. In 2011 the National Science Foundation created a similar list as a way to celebrate their 60th anniversary, and once again PSID was named. PSID has also had a significant influence on cross-national research by inspiring and influencing parallel household panel studies in other countries such as the U.K., (British Household Panel Survey and now the U.K. Household Longitudinal Study), Germany (German Socio-Economic Panel), and Australia (Household, Income and Labour Dynamics in Australia Survey).

Research and publications

Because of its long history and unique design of following adult children as they form their own households, PSID is uniquely positioned to address the next generation of emerging social and behavioral research questions and related policy issues. The breadth of topics that the data archive can be used to address is vast and includes: changes in wealth holdings, housing choices, employment and income, consumption expenditures, family formation and dissolution, health and well-being, and the transition to adulthood; the interplay between economic, social, and health outcomes over the life course and across generations within the same family; the connections between time use and child development, disability, well-being, income, wealth, and consumption expenditures; the effects of early life events on later life economic, health, and psychosocial outcomes; and the changing dynamics of economic status and poverty in particular. The significance of PSID data archive is best summarized by its exceptional record of scholarly productivity. As of September 1, 2018, there were 3,467 published journal articles, 870 dissertations, and 683 books and book chapters based on PSID data, or 5,020 peer-reviewed publications in total. In the most recent year, a peer-reviewed publication that used PSID data appeared in the literature every one and a quarter days. Moreover, PSID-based publications are appearing in the top economic and sociology journals.

During calendar year 2017, there were about five million total hits to PSID website and nearly 220,000 unique visitors. The PSID Data Center provides two ways to download data. The first way is to create customized datasets by selecting various types and years of data and variables. The second way is to download complete data files that are compressed in zip packages. In calendar year 2017, about 16,000 downloads were made directly through the Data Center, and 26,000 zipped downloads were made, for a grand total of about 42,000 data downloads during this period by approximately 3,500 different registered users. Since September 1, 2006, individuals wishing to download PSID data have been required to provide basic information including their email address, name of institution or organization, affiliation (academic, governmental, private, other), and scientific field or discipline. As of the end of calendar year 2017, there were nearly 31,000 registered users, an increase of

nearly 3,000 registered users since the end of 2016. Economics was identified by 72% as their major field, sociology by 9%, with the remainder distributed across education, psychology, demography, child development, medicine, geography and “other.”

Future Directions for PSID

With the collection of the 41st wave of data in 2019, PSID will have followed nearly 5,000 families for 50 years. The study continues to be unique among national data collection efforts in terms of its length and intergenerational design. The study has remained scientifically relevant by adding new content areas and targeted supplements. It has remained scientifically rigorous by maintaining its high response rates and high data quality. And PSID has been at the forefront of data collection methods as well, changing from face to face to telephone and from paper to computer-assisted interviewing as the times demanded.

As we consider the future, we cannot anticipate the breadth of scientific questions that will emerge over the next fifty years. We will continue to look to the scientific community, as we have in the past, for suggestions on salient topics (Moffitt et al. (2015)). In addition, we believe new collection methods are important to integrate into the PSID design in order to provide fresh opportunities to study family life in America. Starting in 2021 we plan to add web as a data collection mode for the core interview along with telephone and face-to-face modes. We have recently developed and tested a self-administered web version of the 2015 telephone instrument, and we are now in the process of updating and expanding the web instrument for testing during 2019–2020 in anticipation of offering web as an option to some families in 2021.

We also plan to continue evaluating the feasibility of between-wave web-based supplements (with a paper questionnaire option), which will provide PSID with flexibility for capturing emerging scientific questions. These between-wave supplements are attractive for gathering information that does not need to be synchronized with the timing of Core (e.g., retrospective information) and can be collected without lengthening the main interview. The mode is also ideal for collecting information directly from more than one family member (rather than relying on one respondent to report for the entire household). A web/mail supplement is also less expensive than telephone-based data collection, which we believe is important given the current constraints on federal and foundation budgets. Web/mail data collection also offers the benefit of quicker collection (the period from design to the end of fieldwork is approximately 6–12 months).

Along with this move to add web as a mode, we are also aware that it may only be possible to collect certain important measures through face-to-face contact. The measures for which face-to-face collection is desirable—or, indeed, essential—includes: collecting anthropometric measurements, physical functioning assessments, cognitive assessments, and observations of the home environment. Biologic material (e.g. saliva) to obtain genetic information can be collected in person or through mail.

In 2019, we will also collect the 8th wave of the Transition into Adulthood study with a web/telephone mixed mode design. Starting in 2017, TAS expanded and harmonized its measures of family formation, including contraceptive use, fertility-related behavior, and cohabitation, with those of the National Survey of Family Growth, a primary U.S. data source for tracking trends in family change through repeated cross-sections (but with no panel component).

Since the original CDS was launched on a cohort of children in PSID families in 1997, the importance of a regular data collection focused on *all* PSID children has become clear. The collection of CDS in 2019 will continue the approach begun in 2014 to conduct the CDS every five-six years on *all* PSID children under age 18 years of age. This strategy ensures that, going forward, information is collected at least three times during childhood for each PSID sample member. We plan to continue our collection of saliva samples from children and primary care-givers. These data will be invaluable for addressing topics such as: identifying the causal effects of measured genotypes on children's development and behavior problems; studying gene-environment interactions that include detailed information on the experience of the family and neighborhood poverty environment; and studying the genetic basis for intergenerational correlations in risky behaviors and developmental outcomes. We are exploring options for expanding the collection of genetic markers for other sample members, including other adult members of CDS households, non-coresident adult relatives such as grandparents, and all other adult relatives of children in CDS. We are also exploring genetic markers beyond genetic sequencing that has recently been undertaken.

At the other end of the life-cycle, the Baby Boom generation, born 1946–1964, were children or young adults when PSID began. The leading edge of this cohort is now in their early 70s. Over the next few decades this generation will transition into the ages where family care demands and other long-term care arrangements are prominent. Because of its long panel design, PSID uniquely offers researchers the ability to examine early and midlife influences on these important later-life outcomes. In addition, its intergenerational design will support study of how family networks respond to life events such as memory impairment and physical limitations.

With the addition of time use to the core instrument in 2017, PSID is the only national study that contains a comprehensive set of economic life-cycle budget components: income, consumption, wealth, and time use. Users will be able to study not only disparities in each of the components but also their distinctive influence on subsequent inequalities in health and wellbeing of older adults as well as implications for human capital formation of the next generation. Because PSID collects this information across generations and over time, it represents the gold standard for evaluating mobility across generations in both economic and health-related outcomes.

The new cohort of immigrants added in 2017 will open a myriad opportunities for scientific discovery. All children from these new immigrant families will also be incorporated into future waves of the ongoing CDS, adding rich information about these children's health, development, and wellbeing. Over time, the sample will support innovative analyses of the lifecourse trajectories of the various immigrant cohorts, including their adaption and

integration across multiple generations, and how various factors have influenced their health and well-being.

Through the years we have linked PSID to a variety of administrative data sets to study, for instance, mortality, medical care utilization, education, housing, and neighborhood effects. PSID is currently exploring many new opportunities for adding low cost linkages to PSID, including housing characteristics from Zillow and Trulia, and related family members from the 1940 Census. The benefits to the project include improving measurement, validation of self-reports, reducing costs by substituting linked data for reported information, and obtaining information not available from respondents (e.g., linked data from historical censuses).

We are investigating the possibility of initiating a joint statistical project with the Census Bureau for the purpose of assigning Protected Identification Keys (PIKs) to PSID respondents. PIKs would allow linkages to internal Census Bureau data sources, including those from the Decennial Census, American Community Survey, the Master Address File data, administrative program data such as food and housing assistance data, other Census data such as from the Longitudinal Employer Household Dynamics (LEHD) program, and possibly Social Security Administration earnings and benefit data. Access to the linked data would be provided through the Federal Statistical Research Data Center (FSRDC) system, with proposal review for access conducted jointly by PSID and Census. Census would be responsible for disclosure review using PSID-defined protocols.

Finally, PSID has made ongoing enhancements to the Data Center and the usability of PSID data. These improvements will continue, including developing new generated variables and Data Center functionalities that will reduce the complexity of the data for new users by facilitating linkages across PSID supplements and across generations of the same family. A more user friendly database that provides the longitudinal data in one dataset is also a possible enhancement.

A Guide to the Remainder of this Volume

The remainder of this volume is dedicated to reviews of key PSID contributions to important social science research topics.

The first article, by Smeeding (this volume), highlights many examples of where PSID has been useful in policy discussions, from intergenerational mobility to long-term welfare dependence. He identifies the degree of equality of opportunity in the U.S. as a key policy issue. He raises the questions: “Do all children, irrespective of their parents’ socioeconomic status, have the same chances of achieving lifetime economic success? Or, are there fundamental disparities that persist from generation to generation creating dynastic patterns of separation in society which will ultimately lead to an even more unequal society?”

The other nine articles demonstrate the importance of longitudinal data by revealing and reviewing how PSID has provided essential information about the well-being and fortunes of multiple generations of Americans. All of the articles highlight the effects of early life

events (e.g., parental income, employment, family background, or access to health care) on well-being in adulthood.

Since its inception, PSID has been used to examine the annual changes in earnings and income (or volatility) and the long-run changes in well-being (or mobility) (see Smith and Morgan (1970)). Here, Moffitt and Zhang (this volume) use PSID to examine changes in short term income volatility. They show that volatility increased from the 1970s to the mid-1980s, leveled off until the early 2000s, and then continued increasing. They highlight how PSID benefits this line of research, because of the length of the panel, the comprehensive measures of socioeconomic status contained in the data, and the rules that allow PSID it to maintain a representative sample.

Mazumder's chapter () highlights the strengths of PSID for intergenerational analysis because it tracks individuals as they form new households with an array of socioeconomic measures over their lifecycles. He reviews many of the studies that have examined intergenerational mobility, which demonstrate that the U.S. is one of the least economically mobile countries among advanced economies. He further compares the PSID results to recent literature that uses tax data to evaluate income mobility and maintains that PSID has advantages over administrative data since it includes many questions that are critical for research purposes but not necessary for administrative purposes.

Leibbrand and Crowder () evaluate mobility in terms of geography, answering the question "where do people move?" They demonstrate that PSID provides insight into geographical mobility by race and over time as individuals grow up. Using information about the specific neighborhoods where children grow up, they highlight the ability of PSID to link family members across generations, and to use geographic specific data to provide context to the places that PSID individuals are moving to and from. Using this geographic rich data they find that neighborhood poverty have different effects on blacks and whites.

One of the key uses of PSID is to examine the life course consequences of childhood experiences and circumstances. How do life events that children experience affect their well-being as adults? Duncan et al () highlight studies that consider the long-term, lifecourse effects of family income during childhood. Key outcomes are earnings and health in adulthood, and their goal is to understand how family circumstances during childhood affect the accumulation of human and health capital and the consequences for later lifecourse outcomes. They argue that revisiting and re-evaluating childhood explanatory variables is well-advised, in light of consistent evidence of the importance of childhood poverty for subsequent life outcomes. From the inception of the work relating childhood poverty to child outcomes, economists, sociologists, and psychologists have offered field-specific theoretical frameworks and empirical tests of those frameworks.

Hofferth et al. () focus on the contributions of PSID Child Development Supplement. They describe research using the CDS to study parental time and monetary inputs to children, the home environment, and the effects of neighborhood context on children's outcomes. They review a substantial body of literature that has used to CDS studies to show how childhood experiences impact development and the physical and mental health of the children as they

grow up. Having a record of behavior over the course of childhood also facilitates the exploration of an alternative pathway, such as family stress or parental emotional problems.

Johnson () uses the ability of PSID to jointly examine the onset of specific health conditions, early childhood experiences, and the characteristics of the neighborhoods throughout the life course. He uses the detailed geocoded data on childhood residential location to show that there are longer-term impacts from access to public health insurance in childhood, on the educational attainment, earnings and health status as an adult. He provides evidence on the influence of childhood neighborhood conditions on the onset of hypertension through middle-age and the long-run impacts of childhood Medicaid spending.

Because of its 50 year collection of data from the same individuals, PSID now has a wealth of information on the other end of the life course – the elderly. Wolf () examines the use of PSID to support research in aging by reviewing a substantial set of research. PSID, with its ability to address individual-and family-level variation at the same time as geographic variation in the service-delivery and programmatic environments, is well-positioned to support research on the consequences of living and care arrangements.

Laditka and Laditka () illustrate the strengths of PSID for providing insights across the lifecourse and generations. They use the PSID to show the salient role that childhood adversity plays in the experience of work disability in adult life. They find that African American, Hispanic, and white women and men with high adversity had shorter lives and more moderate and severe work disability throughout adult life than those reporting no adversities.

Finally, Couch et al. () demonstrate the importance of using PSID to examine lifecourse events from a cross-national perspective. They examine the consequences of union dissolution in the United States on the subsequent economic well-being of partners and children, and compare the results for Germany using the Socio-Economic Panel. The initiation of PSID and the collection of the SOEP also lead to the creation of a data archive containing comparable measures across similar international surveys.

After fifty years of collecting data on the same related individuals and families, PSID looks forward to future decades of data collection and the emerging research questions that can be evaluated with multiple generations across the entire life course for a substantial sample of people. As PSID founder, James Morgan, reflected when asked why PSID was successful, “PSID has lasted because we collect data about real problems and make it available to everyone.” It is this commitment to collecting data about real families and providing the data to researchers who examine emerging problems that will yield a successful PSID for another 50 years.

REFERENCES

- Andreski P, Li G, Samancioglu MZ, & Schoeni R (2014). Estimates of Annual Consumption Expenditures and Its Major Components in the PSID in Comparison to the CE. *The American Economic Review*, 104(5), 132–135. [PubMed: 25221338]
- Cynamon BZ, & Fazzari SM (2015). Household Income, Demand, and Saving: Deriving Macro Data With Micro Data Concepts. *Review of Income and Wealth*.

- Duffy D (2011). 2007 PSID Income and Wage Imputation Methodology. Panel Study of Income Dynamics Technical Paper No. 11–03. Ann Arbor, MI.
- Fitzgerald J, Gottschalk P and Moffitt R. (1998). An analysis of sample attrition in panel data. *The Journal of Human Resources*, 33(2), 251–99. <http://www.jstor.org/stable/146433>
- Fitzgerald J (2011). Attrition in models of intergenerational links in health and economic status in the PSID. *Berkeley Electronic Journal of Economic Analysis & Policy* Vol. 11: Issue 3 (Advances), Article 2.
- Insolera Noura and Freedman Vicki A.. “Comparing Health Estimates in the PSID and NHIS, 2001–2015,” Institute for Social Research, University of Michigan, Technical Series Paper #17–01
- Killewald A, Andreski P, & Schoeni RF (2011). Trends in Item Nonresponse in the PSID, 1968 –2009. (Panel Study of Income Dynamics Technical Series Paper No. 11–02). Survey Research Center, University of Michigan, Ann Arbor, MI.
- Krueger D, Mitman K, & Perri F (2015). Macroeconomics and Household Heterogeneity. *Handbook of Macroeconomics*.
- Lucas Richard, Freedman Vicki A. & Carr Deborah (2018) Measuring experiential well-being among older adults, *The Journal of Positive Psychology*, DOI: 10.1080/17439760.2018.1497686
- McGonagle KA, & Sastry N (2015). Cohort Profile: The Panel Study of Income Dynamics’ Child Development Supplement and Transition into Adulthood Study. *International Journal of Epidemiology*, 44(2), 415–22. [PubMed: 24706732]
- McGonagle Katherine A., Schoeni Robert F., Sastry Narayan, and Freedman Vicki A. The Panel Study of Income Dynamics: Overview, Recent Innovations, and Potential for Life Course Research. *Longitudinal and Life Course Studies*. 2012 3, (2): 268–284. PMID: PMC3591471
- Moffitt RA, Schoeni RF, Brown C, Chase-Lansdale P, Couper M, Diez-Rouz AV, Hurst E, & Seltzer JA (2015). Assessing the Need for a New Nationally Representative Household Panel Survey in the United States. *Journal of Economic and Social Measurement*, 40, 1–26. [PubMed: 26688609]
- National Research Council. (2013). *Nonresponse in Social Surveys: A Research Agenda*. Tourangeau R & Plewes TJ (Eds). Washington DC, National Academies Press.
- Office of Economic Opportunity, “Third Annual Report of the Office of Economic Opportunity,” Government Printing Office, 1968
- Panel Study of Income Dynamics Main Interview User Manual. (2017). Institute for Social Research, University of Michigan, May, 2017.
- Pfeffer FT, Schoeni RF, Kennickell A, & Andreski P (2014). Measuring Wealth and Wealth Inequality: Comparing Two U.S. Surveys. Panel Study of Income Dynamics. Technical Series Paper No. 14–03. Survey Research Center, University of Michigan, Ann Arbor, MI.
- Sastry Narayan, Fomby Paula, and McGonagle. Katherine “Using the Panel Study of Income Dynamics (PSID) to conduct life course health development analysis” In *Handbook of Life Course Health Development*, pp. 579–599. Springer, Cham, 2018.
- Schoeni Robert F. and Wiemers Emily. The Implications of Selective Attrition for Estimates of Intergenerational Elasticity of Family Income. *The Journal of Economic Inequality*. 2015 13, (3): 351–372. [PubMed: 26251655]
- Sirken MG (1970). Household surveys with multiplicity. *Journal of the American Statistical Association*, 65(329), 257–266.
- Smith James and Morgan James N. Dynamics of Income Distribution: Poverty and Progress: Variability of Economic Well-being and Its Determinants. *American Economic Review*. 1970 60, (2): 286–295.