Departments

In this issue—

• Data Shop
Data Shop

Data Shop, a department of Cityscape, presents short articles or notes on the uses of data in housing and urban research. Through this department, the Office of Policy Development and Research introduces readers to new and overlooked data sources and to improved techniques in using well-known data. The emphasis is on sources and methods that analysts can use in their own work. Researchers often run into knotty data problems involving data interpretation or manipulation that must be solved before a project can proceed, but they seldom get to focus in detail on the solutions to such problems. If you have an idea for an applied, data-centric note of no more than 3,000 words, please send a one-paragraph abstract to david.a.vandenbroucke@hud.gov for consideration.

Using the Panel Study of Income Dynamics To Analyze Housing Decisions, Dynamics, and Effects

Katherine McGonagle
Narayan Sastry
University of Michigan

Abstract

The Panel Study of Income Dynamics (PSID) is the world's longest running household panel survey. It started in 1968 and has followed the same families—and their descendants—for nearly 50 years. PSID was conducted annually from 1968 through 1997 and has been conducted biennially since 1997. As of 2015, 39 waves of data have been collected. In 2015, interviews were completed with more than 9,000 households and information was collected on about 25,000 household members. PSID has achieved high wave-to-wave response rates throughout most of its history. Since the beginning of the study, detailed information has been collected on family composition, income, assets and debt, public program participation, and housing. At the beginning of the recent housing crisis, PSID began collecting information about mortgage distress and foreclosure activity. PSID currently includes several major supplemental studies. The Child Development Supplement and the Transition into Adulthood Supplement collect detailed information about behavior and outcomes among children and young adults in PSID families, such as educational achievement, health, time use, family formation, and housing-related decisions among young adults. PSID data are publicly available free of charge.
Abstract (continued)

to researchers; some data available only under contract to qualified researchers allow linkage with various administrative databases and include information such as census tract and block of residence that can be used to describe neighborhood characteristics. PSID data have been widely used to study topics of major interest to Cityscape readers, including housing decisionmaking, housing expenditures and financing, residential mobility and migration, and the effects of neighborhood characteristics on a variety of measures of child and family well-being. This article provides an overview of PSID and its housing- and neighborhood-related measures. We briefly describe studies using PSID on housing-related topics. Finally, we point readers to resources needed to begin working with PSID data.

The Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID) is the world’s longest running, nationally representative household panel study, with information collected on sampled families and their descendants for nearly 50 years. PSID began in 1968 to gauge the success of President Lyndon Johnson’s “War on Poverty” and to track the economic well-being of U.S. families. Housing and neighborhood characteristics are key indicators of family economic well-being and have been included in the study since its inception.

PSID began with a national sample of about 5,000 households with approximately 18,000 individuals (Hill, 1992). The study has followed these individuals and their descendants at each wave, leading to sample growth over time. PSID’s 2015 wave includes about 10,000 households containing 25,000 individuals. Respondents have been interviewed by telephone since 1973, with interviews conducted annually from 1968 to 1997 and biennially thereafter. Wave-to-wave core reinterview response rates typically range between 96 and 98 percent. PSID data are available free of charge to the public and have been used for approximately 4,000 peer-reviewed publications, including more than 700 dissertations. The study’s design has been replicated in many countries around the world. PSID is regularly used for policy analysis by U.S. federal government agencies. On the National Science Foundation’s (NSF’s) 60th anniversary, it named PSID as 1 of the 60 most significant scientific advances ever funded by NSF.

PSID’s unique features include its national representativeness, the long duration of the panel, its genealogical design, and its broad and deep content. PSID includes adult respondents of all ages and follows individuals across the entire lifecourse. Adult children are interviewed in their own family units after they achieve economic independence from their parents’ households. This unique self-replacing design means that, for many families, PSID includes self-reported information on three (and occasionally four, or even five) generations of the same family at various points in their lifecourse. PSID is the only survey ever collected on lifecourse and multigenerational economic conditions in a long-term panel representative of the full U.S. population (see McGonagle et al.,
Using the Panel Study of Income Dynamics To Analyze Housing Decisions, Dynamics, and Effects

Information on Housing and Neighborhoods

Considerable information about housing and neighborhood characteristics has been collected in every wave of PSID (see exhibit 1). Topics include dwelling characteristics, housing utilities, residential mobility and migration, housing-related financial information and consumption expenditures, mortgage distress, and neighborhood characteristics based on geospatial identifiers and administrative data. Information on home and neighborhood characteristics and the emergence of financial independence and housing-related decisionmaking has also been collected in CDS and TAS.

Since the start of PSID, data have been collected on dwellings characteristics, including dwelling type and number of rooms. Information is also collected about characteristics of retirement and senior

## Exhibit 1

### Housing and Neighborhood-Related Questionnaire Content in PSID-CDS-TAS

<table>
<thead>
<tr>
<th>PSID Housing-Related Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dwelling characteristics</strong></td>
</tr>
<tr>
<td>Housing type: house, duplex, apartment, condominium, townhouse, mobile home.</td>
</tr>
<tr>
<td>Number of rooms.</td>
</tr>
<tr>
<td>Number of individuals sharing living space.</td>
</tr>
<tr>
<td>Whether retirement or senior community and services offered.</td>
</tr>
<tr>
<td><strong>Housing utilities: type, cost, and frequency of payments</strong></td>
</tr>
<tr>
<td>Home heat, water, sewer, electricity, telephone service, air-conditioning, cable television, Internet connection.</td>
</tr>
<tr>
<td>Use of government programs for utility costs.</td>
</tr>
<tr>
<td><strong>Housing finances</strong></td>
</tr>
<tr>
<td>Whether owns or rents.</td>
</tr>
<tr>
<td>Current market value of dwelling.</td>
</tr>
<tr>
<td>Rental agreement detail.</td>
</tr>
<tr>
<td>Mortgage detail.</td>
</tr>
<tr>
<td>Property tax amounts.</td>
</tr>
<tr>
<td>Home insurance amounts.</td>
</tr>
<tr>
<td>Use of government programs for housing.</td>
</tr>
<tr>
<td><strong>Housing consumption expenditures</strong></td>
</tr>
<tr>
<td>Annual expenditures for home repairs and maintenance and for household furnishings and equipment.</td>
</tr>
<tr>
<td><strong>Mortgage distress</strong></td>
</tr>
<tr>
<td>Falling behind in housing payments.</td>
</tr>
<tr>
<td>Foreclosure activity.</td>
</tr>
<tr>
<td>Mortgage modifications.</td>
</tr>
<tr>
<td>Second mortgages.</td>
</tr>
<tr>
<td>Expectations about housing payment difficulties in coming year.</td>
</tr>
<tr>
<td><strong>Residential mobility, reasons for moving, moving intentions</strong></td>
</tr>
<tr>
<td>Residential change timeline, including timing and address of all residential moves occurring during past 2 calendar years.</td>
</tr>
<tr>
<td>Reason for moving, including changes in employment, school attendance, or marital status; for an improved living situation; to save money; or because of a financial shock, such as bankruptcy, foreclosure, or eviction.</td>
</tr>
<tr>
<td>Likelihood of moving in near future and reason.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CDS and TAS Housing and Neighborhood-Related Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neighborhood characteristics and home environment (CDS)</strong></td>
</tr>
<tr>
<td>Ratings by primary caregiver on neighborhood stability, social cohesion, safety, and satisfaction.</td>
</tr>
<tr>
<td>Ratings by interviewer on HOME Inventory, including availability of reading material, technology, musical instruments, and toys; features of play areas; lighting adequacy; clutter; cleanliness; space; noise; and condition of nearby homes and buildings.</td>
</tr>
<tr>
<td><strong>Location and economic independence (TAS)</strong></td>
</tr>
<tr>
<td>Where living during different parts of year, including parent’s home, college dormitory, apartment or rented home, military base, or other institution.</td>
</tr>
<tr>
<td>Whether moved for an employment opportunity.</td>
</tr>
<tr>
<td>Help received from parents and relatives for housing payments and amounts received.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Restricted Use Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assisted housing administrative linkages (PSID)</strong></td>
</tr>
<tr>
<td>Type and class of subsidy, including public housing, low-income housing tax credit, Farmers Home Administration, other federal, other state, other project-based housing, other tenant-based housing.</td>
</tr>
<tr>
<td><strong>Geospatial data (PSID-CDS-TAS)</strong></td>
</tr>
<tr>
<td>For all waves: FIPS county and place; 5-digit ZIP Code; MSA and CBSA; census tract, block, and block group; match-quality indicators.</td>
</tr>
</tbody>
</table>

---

housing communities. Since the earliest waves, information has been collected on the type and cost of utilities, including source of home heat, air-conditioning, and the use of government subsidies for utility costs. Questions were added more recently about cable television and Internet connections.

Data on residential mobility, moving intentions, and reasons for moving have been collected throughout the study. For each wave, information is obtained on all recent residential moves and their timing. Specific reasons for each move are also collected. Respondents provide estimates of the likelihood of moving in the near future and describe life events that may trigger potential moves.

Detailed housing finance information has been obtained since the earliest waves of PSID, including current market value of the dwelling, details of rental agreements and mortgages, and the use of government subsidies. Starting in 2005, data have been collected on housing-related consumption expenditures, including annual costs of home repairs and maintenance and of household furnishings and equipment as part of a complete series on consumption expenditures.

At the onset of the 2009 housing crisis, PSID began collecting extensive information about mortgage distress, including falling behind in payments, mortgage modifications, foreclosure activity, and expectations about housing payment difficulties in the coming year. This information has been used extensively to describe and analyze families' difficulties with home mortgages during the Great Recession (December 2007 to June 2009), including by the Federal Reserve Board (for example, Sherlund, 2010) and by others (for example, Lin, Liu, and Xie, 2016).

Although nearly all the data are freely available in the public domain, certain information about housing and geography is available only through a restricted data use contract to maintain the confidentiality of PSID respondents, including geospatial identifiers below the level of state and administrative linkages to external databases. These geospatial identifiers and administrative data have been widely used as a means of characterizing the neighborhoods in which respondents live. Three levels of geospatial data are available: census tract, block group, and block. Residential addresses have been geocoded for all waves of the study using four different versions of census geography: addresses from 1968 through 1985 were geocoded using both the 1970 and 1980 census geography; those from 1968 through 1999, using the 1990 census geography; those from 2001 through 2009, using the 2000 census geography; and those since 2011, using the 2010 census geography.

Linkages of PSID data to administrative records, including the receipt of government housing subsidies, are also available. These linkages are generated through a process that matches addresses of PSID families in each wave with those in the Assisted Housing Database collected by the U.S. Department of Housing and Urban Development (HUD). Information is available regarding whether a PSID address in a given year corresponds to an assisted housing address, and, if so, the type of assisted housing, including whether subsidized by HUD, by the former Farmers Home Administration, by tax credits administered by the U.S. Department of the Treasury, or through state-level housing subsidy programs.

Other administrative data include identifiers for primary and secondary schools attended by children in CDS and TAS. These school identifiers link PSID children to detailed information about their schools from the Common Core of Data and Private School Universe Survey prepared by the U.S. Department of Education's National Center for Education Statistics (NCES). PSID and TAS
sample members who have attended college, university, or technical and vocational postsecondary institutions have identifiers that can be linked to data from the NCES Integrated Postsecondary Education Data System.

With a central goal of providing information about how child health and well-being are influenced by home and neighborhood environments, CDS has collected particularly rich information on these latter topics. All waves include detailed information collected from the child’s primary caregiver on neighborhood stability, social cohesion, safety, and satisfaction. The Home Observation Measurement of the Environment Inventory (Caldwell and Bradley, 2003), designed to measure the quality and quantity of stimulation and support available to a child in the home environment, has been included in all waves of CDS.

Finally, TAS collects housing-related content that reflects the high rates of mobility and emerging financial independence characteristic during young adulthood. For instance, information is obtained about where young adults live “most of the time,” including in parent’s home, a college dormitory, an apartment or rented home, a military base, or other institution. Because young adults move frequently, this information is collected for different parts of the year (October through April and May through August); information regarding whether a move occurs for an employment opportunity is also captured. TAS also assesses young adults’ economic independence by collecting information about help received from parents and relatives in paying rent or a mortgage. Additional data are collected across many other domains, such as self-perceptions; future expectations for schooling, careers, and employment; and information regarding health, wealth, and income that can support rich models of housing decisions and their effects on social and economic outcomes during young adulthood.

Studies Using PSID on Housing-Related Topics

Data collected in PSID, CDS, and TAS have supported a large body of scientific work across a variety of topics related to housing and neighborhood characteristics. A comprehensive bibliography of PSID publications is available on the project’s website.

Assisted Housing

A number of studies have used the PSID Assisted Housing Database to examine the consequences of receiving subsidized housing. Newman and Harkness (2000) found that the lower educational attainment of children who lived in public housing disappeared once measured characteristics were taken into account. In another set of analyses, also exploiting PSID’s longitudinal design, Newman and colleagues have examined the effects of housing assistance on employment outcomes and welfare receipt (Harkness and Newman, 2003; Newman and Harkness, 2002; Newman, Holupka, and Harkness, 2009). This research shows no negative effects on employment outcomes, although public program participation rates are higher in the future. In a paper that exploits the intergenerational richness of PSID, Kucheva (2014) found that adults who grew up in subsidized housing had a higher probability of residing in subsidized housing in adulthood.
**Neighborhood and Housing Choice**

PSID provides a rich data source for examining choices about neighborhood and housing choices. A number of studies examined the dynamics of housing tenure choices by families, examining transitions between homeownership and rental tenure and the factors associated with these transitions (Bajari et al., 2013; Boehm and Schmittmann, 2014; Börsch-Supan and Pollakowski, 1990; Carter, 2011; Henderson and Ioannides 1989; Ioannides, 1987; Kan, 2000). Ties between housing and neighborhood choice were examined using PSID data, focusing, for instance, on the process of “downsizing” of housing and retirement moves among the elderly (Banks et al., 2012; Bian, forthcoming; Painter and Lee 2009; Sabia, 2008; VanderHart, 1998). PSID was used to examine the effects of neighborhood characteristics on housing decisions (for example, Lee, 2014) and also the consequences of individuals’ residential decisions on neighborhood dynamics (for example, Bruch, 2014).

**Effects of Neighborhood Characteristics**

PSID has been used extensively to investigate the effects of neighborhoods, as evidenced by hundreds of publications on this topic. PSID was one of the earliest data sources for studying contextual effects on socioeconomic status (Corcoran et al., 1990; Dachter, 1982) and remains one of the most important and widely used sources across multiple disciplines for examining neighborhood effects on a variety of outcomes, including child, adolescent, and young adult development (Dearing et al., 2009; Jackson and Mare, 2007; Sastry, 2012; Sharkey and Elwert, 2011; Timberlake, 2009a, 2009b; Wimer et al., 2008); health (Do and Finch, 2008; Do, Wang, and Elliott, 2013; Halliday, 2007; Halliday and Kimmitt, 2008; Johnson, 2012; Wen and Shenassa, 2012); education (Brooks-Gunn et al., 1993; Crowder and South, 2011, 2003; Galster et al., 2013, 2007; Harding, 2003; Wodtke, Harding, and Elwert, 2011); income and earnings (Islam, 2013; Sharkey, 2012, 2008); the intergenerational transmission of neighborhood context (Dawkins, 2005a; Sharkey, 2008; Sharkey and Elwert, 2011; Solon, Page, and Duncan, 2000); family migration and labor force outcomes (Blackburn, 2010; Shauman, 2010; Shauman and Noonan, 2007; Swain and Garaskey, 2007); and fertility behavior (Clark and Withers, 2009; South, 2001a, 2001b; South and Crowder, 2011, 1999; Wodtke, 2013). With an oversample of African-American families, PSID is a key data source for examining levels and trends in residential segregation by race (Crowder and Downey, 2010; Crowder and South, 2005; Dawkins, 2005b, 2006; Freeman, 2008, 2005a, 2005b; Pais, South, and Crowder, 2012; Sharkey, 2012, 2008; South and Crowder, 2005; South, Crowder, and Pais, 2011; Timberlake, 2007; Vartanian, Buck, and Gleason, 2007; Wagmiller, 2013; White et al., 2005).

There are many opportunities for new research on the effects of neighborhood characteristics. In particular, the continued collection of data in PSID and new data from CDS will support new studies that build on previous research by Crowder and South (2011), Harding (2003), Wodtke, Elwert, and Harding (2012), Wodtke, Harding, and Elwert (2011), and others who used PSID to examine contextual effects on high school graduation and found important effects of neighborhood concentrated disadvantage. The information obtained from the new cohort of children in CDS and young adults participating in TAS will enable researchers to examine how health, development, and well-being today are shaped by several key features of parents’ and grandparents’ past.
environments—especially the consequences of growing up in poor neighborhoods. PSID has collected unparalleled nationally representative data every 1 or 2 years during the past four decades that enable researchers to accurately characterize, using contemporaneous measures, children’s, parents’, and grandparents’ experiences of growing up in a poor family and in a poor neighborhood. As a result, PSID and its supplemental data on children and young adults provide essential information for studying the replication of poverty and advantage across generations and the life-course. Further, with the rich data on the home, neighborhood, and school environments available today, researchers can examine the pathways through which developmental outcomes are affected by poverty and socioeconomic status. Results of these analyses will provide valuable information for policymakers to improve the lives of disadvantaged children in the United States.

Effects of the Great Recession and Housing Crisis

Research to date using PSID has described the direct economic consequences of the Great Recession and associated housing crisis on wealth, job losses, consumption expenditures, and retirement decisions (for example, Attanasio and Pistaferri, 2014; Bosworth, 2012; Parent, 2015; Pfeffer, Danziger, and Schoeni, 2013); residential mobility (Coulson and Grieco, 2013); charitable giving (Marx and Carter, 2014); and household formation (Lee and Painter, 2013). Other work has used PSID data to describe foreclosure risk for individual households and disparities in this risk by race and ethnicity (for example, Hall, Crowder, and Spring, 2015).

PSID data can be used to study how the economic effects of the Great Recession and housing crisis translate into lifecourse decisions about schooling, employment, and residential preferences and consequences for educational attainment, health, and well-being. For example, recent work shows that change in a household’s housing wealth in the 4 years prior to a child being of college age reduces the likelihood that the child will attend college (Lovenheim, 2011). The ongoing data collected through PSID and TAS provide an unprecedented opportunity to examine how these national financial adversities, combined with secular changes in federal financial and mortgage policies, will ultimately shape residential preferences of young adults. Moreover, recent data collected from children in the new CDS-2014 were drawn from a population that lived through the Great Recession and that experienced higher levels of parental unemployment and poverty than during any time since the early 1990s (Isaacs, 2011). The circumstances of these children can be compared with a previous generation of children who participated in the original CDS from before the financial crisis to study questions such as the impact of the housing and foreclosure crisis on outcomes such as child behavioral problems through family experiences or neighborhood exposures.

How To Access the Data

Most PSID data and documentation are freely and publicly available on the PSID website (http://www.psidonline.org). Information is currently available on more than 70,000 variables, on nearly 75,000 individuals, and for all waves of the PSID and its supplements. Users can create customized data extracts from any set of waves by searching or browsing for variables, can obtain customized codebooks specific to their data extract, and can archive data extracts for shared and future use. They can “load” their data carts with variables by wave. They can view variable descriptions,
including univariate statistics and names of the same variable in other waves, by clicking an “open-book” icon next to each variable. They can edit their cart by removing or adding variables through a return to the “data aisle.” Users may save data carts, enabling them to share specific extracts with colleagues, reviewers, and students. A range of file formats is available when the user is ready to “check out,” including SAS, STATA, SPSS, dBase, Excel, and ASCII. The PSID website provides a cross-year variable index that facilitates searching and browsing all variables across the full archive from 1968 to the most recent wave and for all waves of CDS and TAS. Organized by content domains, the index is integrated with the online Data Center so that users can view the codebook and add variables directly to their data cart from the index. Geospatial data below the level of state and linked administrative data may be obtained after establishing a data use agreement between a user’s institution and the University of Michigan.2 PSID has also made available a set of user tutorials and webinars on a variety of topics, including an introduction to the PSID for the new user3 and provides a Help Desk that gives rapid responses to users’ questions.

**Acknowledgments**

The authors gratefully acknowledge funding support from the National Science Foundation (1157698), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (HD069609, HD059779, HD072493, HD052646), and the National Institute on Aging (AG040213).

**Authors**

Katherine McGonagle is a research scientist in the Survey Research Center at the University of Michigan’s Institute for Social Research and is the Assistant Director of the Panel Study of Income Dynamics.

Narayan Sastry is a research professor in the Survey Research Center and the Population Studies Center at the University of Michigan’s Institute for Social Research and is an associate director of the Panel Study of Income Dynamics.

**References**


2 Information about PSID restricted data may be found at the PSID Data center (http://www.psidonline.org/).

3 Tutorials available on the PSID website explain the content of the survey, how to access the data, and how to use the PSID online Data Center; see http://psidonline.isr.umich.edu/Guide/tutorials/default.aspx.


Using the Panel Study of Income Dynamics To Analyze Housing Decisions, Dynamics, and Effects


