

## EXECUTIVE SUMMARY

The lack of quality data has been a long-standing concern among researchers studying assets. Except for the 1962 Survey of Financial Characteristics of Consumers, no serious efforts were made to collect reliable asset data before the 1980s. Beginning in 1983 and 1984, the Survey of Consumer Finances (SCF), Survey of Income and Program Participation (SIPP), and the Panel Study of Income Dynamics (PSID) began collecting asset and liability data. With available data, researchers started to examine asset distribution, test theoretical models and hypotheses, and develop new concepts and theories on assets.

To guide future research in assets, this report examines the following questions:

- What are the most informative and reliable data sources for understanding low-income households' assets and liabilities?
- What are the limitations of these data sources?
- What are the means for improving asset data?

### **Data Sets Identified and Criteria for Assessment**

Our review of the literature, survey data, and demonstration data has identified 12 data sets that have the potential to provide important information on low-income households' assets and liabilities. The 12 data sets are:

1. American Dream Demonstration Account Monitoring (ADD-AM) Data
2. American Dream Demonstration Experiment (ADD-E) Data
3. Assets For Independence Act (AFIA) Evaluation Data
4. Consumer Expenditure Survey (CEX)
5. Current Population Survey (CPS)
6. Health and Retirement Study (HRS)
7. Home Mortgage Disclosure Act (HMDA) Data
8. National Longitudinal Study of Youth 1979 (NLSY79)
9. National Survey of Family and Households (NSFH)
10. Panel Study of Income Dynamics (PSID)
11. Survey of Consumer Finances (SCF)
12. Survey of Income and Program Participation (SIPP)

We evaluate these data sets with four criteria: relevancy, representativeness, recurrence, and richness of correlates. Although we pay special attention to each data set's ability to provide information on assets (relevancy) among low-income population (representativeness) over time (recurrence), we also consider the correlates data sets provide (richness of correlates) and thus their ability to answer other important research questions, such as the effects of assets on outcomes. Exhibit 2 provides a brief overview of the evaluation criteria attained by each data set.

Based on these four criteria, we identify three primary data sets as having the greatest potential for asset research: the Survey of Consumer Finances (SCF), the Survey of Income and Program Participation (SIPP), and the Panel Study of Income Dynamics (PSID). A summary of these three primary data sets is provided in Exhibit 5 and a summary of nine secondary data sets is provided in Appendix A. Below we highlight the strengths and weaknesses of our primary data sets.

## **Primary Data Sources for Holdings of Low-Resource Households' Assets**

### ***Survey of Income and Program Participation***

- The SIPP contains detailed information on financial assets (e.g., checking and saving accounts, stocks, IRAs) and nonfinancial, tangible assets (e.g., personal residences and vehicles), as well as information on households' liabilities (e.g., home mortgages and credit card debt).
- The SIPP has a large sample size that allows researchers to examine detailed subpopulations. The 2001 SIPP panel includes 35,100 households, which is a much larger sample than in both the PSID and SCF.
- SIPP panels can be matched with administrative data containing individuals' earnings histories and employer information. These matched data provide information on individuals' history of contributions to retirement savings accounts (e.g., 401(k) accounts).
- Analyses of the 1990–93 SIPP panels and the SCF show that the SIPP and SCF data sets have similar wealth distributions up through the 80<sup>th</sup> percentile; it is only above the 80<sup>th</sup> percentile that wealth on the SIPP falls significantly short of the SCF (Rodgers and Smith 2000). A comparison of the 1996 SIPP to the 1998 SCF, however, finds that SIPP net worth falls short of SCF net worth across the entire distribution (Czajka et al. 2003).
- The SIPP has relatively high rates of item nonresponse for asset and liability questions, higher than both the PSID and SCF.
- SIPP panels are longitudinal and have typically lasted three to four years. This feature allows researchers to observe changes in asset holdings over time.
- Although the SIPP is currently undergoing a process of review and reengineering, a 2008 panel is planned.

### ***Panel Study of Income Dynamics***

- The PSID asset data provide a good accounting for the major components of net worth and include information on home value, financial assets, tangible assets, and debt.
- The PSID follows the same families longitudinally over several decades, allowing for intergenerational analysis, trend studies, and examination of the influence of family histories.
- The PSID has little missing data for its asset questions.

- The PSID includes a series of “active savings” questions that trace flows of money in and out of assets, such as when a house is bought or sold, money is put into the stock market, or an annuity is cashed in. This allows researchers to study whether changes in net worth come from savings or capital gains.
- The PSID asks questions about significantly fewer assets and liabilities categories than either the SIPP or SCF. The PSID has 9 asset and liability questions covering 12 topic areas as compared to the 100 questions pertaining to 30 assets and liabilities topic areas covered by the SCF.
- Comparing wealth distributions, the PSID is almost identical to the SCF up to the 30<sup>th</sup> percentile, then PSID estimates become a little lower. By the upper 1 to 3 percentiles of the distribution, the two surveys begin to diverge more dramatically, where SCF estimates are higher.
- The PSID has a relatively small sample size (7,823 families in 2003). As a result, many specialized subgroup analyses are not possible (Kim and Stafford 2000).

### ***Survey of Consumer Finances***

- The SCF is the most focused survey on asset ownership, providing the most detailed questions about assets and liabilities.
- The SCF makes a concerted effort to provide an accurate estimate of aggregate net worth in the United States through its sampling procedures—oversampling the wealthy who own a large proportion of the nation’s assets. It also has sophisticated imputation procedures.
- The SCF is a cross-sectional survey that provides asset and liability data every three years, while, in recent years, the SIPP provides this data annually and the PSID biennially.
- The SCF has a relatively small sample size that may make it difficult to study low-income subpopulations. For instance, the 2004 SCF includes only 4,522 households.
- The correlates in the SCF may limit its usefulness for many asset-related studies. For example, the survey does not provide information on the well-being or education level of children in the household.

### **Means for Improving Asset Data**

Our evaluation identifies multiple options for improving the scope and quality of asset and liability data. We examine both general and specific means for improving data on assets. General options are based on limitations common to existing data sets. We also provide options for improving each of three primary data sets. More details about these options for improvement are in the main body of the report.

#### ***General Means for Improving Asset Data***

- ***Option 1:*** Collect information on the dynamics of asset accumulation to better understand how assets change over the life course.

- **Option 2:** Encourage survey respondents to use financial statements when answering asset and liability questions, which will ensure more accurate responses.
- **Option 3:** Assess the quality of asset data using other data sources, particularly for difficult-to-value assets, to provide more consistent valuations.
- **Option 4:** Collect data on assets among subpopulations that may be of interest in the general poverty policy and research arena (e.g., formerly incarcerated individuals, Native Americans, immigrants).
- **Option 5:** Collect data on respondents' experience with saving incentive programs beyond retirement savings accounts to provide a better understanding of saving patterns and responses to savings incentives.

### ***Specific Means for Improving Asset Data***

#### *Survey of Income and Program Participation*

- **Option 6:** Collect information on quasi-liquid pensions, life insurance, other assets (jewelry, cemetery plots, art, collections), and other secured debt, which will provide a comprehensiveness of assets and liabilities accounting equal to that of the SCF.
- **Option 7:** Revise the imputation method used to fill in for item nonresponse to provide more accurate estimates when data are missing.
- **Option 8:** Increase the number of response brackets, narrow the width of each bracket, and raise the top code of asset brackets to allow respondents to give more accurate responses when they are unable to give an exact value of their asset (or liability).
- **Option 9:** Collect data on legal (citizenship and refugee) status more than once during the SIPP panel to allow for examination of how assets and liabilities change with immigrant status.

#### *Panel Study of Income Dynamics*

- **Option 10:** Link PSID data with administrative data (e.g., data from the Social Security Administration and Internal Revenue Service) to obtain information on individuals' earnings histories, contributions to retirement savings accounts (e.g., 401(k) and 403(b) accounts), and contributions to the Social Security system, as the SIPP and HRS currently do.
- **Option 11:** Add questions about business equity and other assets to the wealth section, which would strengthen PSID's ability to capture net worth across the entire wealth distribution.

*Survey of Consumer Finances*

- **Option 12:** Link SCF data with administrative data (e.g., data from the Social Security Administration, Internal Revenue Service, and Employer Pension Study).
- **Option 13:** Add more low-income households to the SCF sample to allow for analyses of asset holdings of low-income subgroups such as racial groups or high school dropouts.
- **Option 14:** Decrease the lower limit of *inter vivos* transfer in the SCF to allow for more accurate examination of inter-family transfers and asset accumulation for low-income households.
- **Option 15:** Collect data as a panel that includes at least two or three waves (i.e., points of data collection), which would provide the information needed to study portfolio changes.