The Measurement of Household Wealth using Survey Data: 
An Overview of the Survey of Consumer Finances

Survey measurement of household wealth faces some inherent difficulties: respondents may not provide financial information, values of assets and liabilities can be hard to report accurately, and ordinary random samples contain too few wealthy households to represent the full distribution of wealth. The Survey of Consumer Finances has a number of special features intended to promote accurate estimation of household wealth. This paper provides an overview of the SCF, including the sample design, response rates, information collected, and data quality and imputation.

Gerhard Fries, Federal Reserve Board
Martha Starr-McCluer, Federal Reserve Board
Annika E. Sundén, Federal Reserve Board

Introduction

To understand the economic circumstances of the population, it is important to collect information on household wealth. Wealth information indicates the extent to which households have savings to draw on in the event of unemployment or illness. It sheds light on the question of how well prepared today's working households will be to finance consumption during retirement. It provides indications of households' ability to service their debts, including their potential vulnerability to default or bankruptcy. It is also valuable for measuring inequality in the distribution of wealth, and for understanding its causes.

There are several reasons why the measurement of wealth through a household survey is inherently difficult. First, compared to other survey topics, respondents may be relatively reluctant to answer questions about wealth and income. Even when assurances of confidentiality are given, respondents sometimes have lingering concerns about the legitimacy of the survey or about the extent of its commitment to preventing disclosure. Second, values of assets and liabilities can be hard to report accurately. For example, respondents may not know the current market value of some of their assets, or they may forget to report some of their assets or debts. Finally, because the ownership of wealth is relatively concentrated, ordinary random samples are not likely to contain enough wealthy households to provide good representation of the full distribution of wealth.

The main U.S. survey collecting information on household wealth is the Survey of Consumer Finances (SCF). This paper provides an overview of the survey's content and methodology. The next section describes the various surveys conducted since the early 1960s. Section 3 discusses the sample design and response rates of recent SCFs. Section 4 summarizes the information collected in the survey. Section 5 covers data quality and imputation. Section 6 reviews the use of the SCF data in research on saving, wealth and other topics.

Background

Surveys of consumer finances have been conducted periodically since the early 1960s, under the sponsorship of the Federal Reserve Board of Governors. While the surveys have differed somewhat in their sample designs and questionnaires, they have shared the objective of collecting detailed information on assets and liabilities from a nationally representative sample of U.S. households. Because the surveys have been intended to provide accurate descriptions of the current financial situations of consumers, the data are primarily cross-section in nature, although some panel data also have been collected.

The various surveys have differed somewhat in their samples and methodologies, so it is useful to summarize their main features (see table 1, which also provides a general reference for each survey). The first survey, called the Survey of Financial Characteristics of Consumers, collected information on consumers' financial situations in 1962. A follow-up interview collected information on changes in consumer finances between 1962 and 1963. Another wealth survey was not conducted again until the 1983 Survey of Consumer Finances, when 4,103 households were interviewed. Some 2,822 of these households were reinterviewed by phone in 1986, using a shorter questionnaire. In 1989, the survey interviewed both new respondents and some respondents from the 1983 SCF. Two separate data sets were produced from the 1989 survey: a cross-section data set intended to provide accurate representation of the population as a whole, and a panel data set consisting of the 1,479 respondents to the
1983 SCF who were reinterviewed in 1989. The questionnaire used in 1989 was similar to that used in 1983, but with considerable revisions in some sections. Since 1992, the SCF samples have been cross-sections only, and the questionnaires have been very similar to that used in 1989, facilitating comparisons of the data across years.

Table 1
Summary of the Surveys of Consumer Finances

<table>
<thead>
<tr>
<th>Survey</th>
<th>Sample size</th>
<th>General Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>List</td>
</tr>
<tr>
<td>1962 SFCC</td>
<td>2,558</td>
<td>-</td>
</tr>
<tr>
<td>1962-63 SCFF</td>
<td>2,165</td>
<td>-</td>
</tr>
<tr>
<td>1989 Cross-section</td>
<td>3,143</td>
<td>866</td>
</tr>
</tbody>
</table>

Note: The 1998 SCF is scheduled to be fielded in the second half of 1998.

All of the surveys have shared two features that make them quite different from other surveys collecting information on wealth. First, because a large share of wealth is held by a relatively small share of the population, the SCFs have taken measures to ensure adequate representation of wealthy households. All of the surveys have used dual sample frames, with a standard representative sample supplemented by a special sample of high-income taxpayers (see below). Second, the survey collects very detailed information on assets and liabilities. In a questionnaire that goes item by item through a list of assets and liabilities, the risk of forgetting to report an item is smaller than with a questionnaire referring to broad categories of assets or liabilities. This approach reduces tendencies toward underreporting, and also provides valuable information on the composition of household wealth.

Sample Design and Response Rates for Recent SCFs

Since 1989, the SCF has had a dual-frame sample design structured as follows. The first part consists of a standard multi-stage area probability (AP) sample, drawn from a file based on Census records. At the first stage, the U.S. is divided into geographical units, which are stratified by population size, urban status, and region; then units are selected into the sample in a way that ensures national representation. At the next stage, smaller areas are selected, and a sample of dwellings is drawn. The main families living in these dwellings are the potential AP respondents. The second part of the sample is a "list sample" drawn from a special file of tax records. Information on asset income is used to compute a proxy for net worth; this proxy is used to stratify the file, and units are selected into the sample at disproportionate rates, with the sampling rate higher for the wealthier strata. Together, the two parts of the SCF sample are intended to provide good coverage of both broadly-distributed items, like homeownership, and items that are more narrowly held, like businesses and bonds.

In descriptive analysis, sample weights can be used to make the data representative of the population as a whole. The weights take into account both the complex sample design (for example, cases from high-wealth strata generally receive very small weights) and nonrandom aspects of nonresponse. In recent SCFs, the response rate has been around 70 percent for the area-probability sample, which is not unusual for a long cross-section survey. For the list sample, the share of cases eventually responding to the survey varies from 43 percent among moderately wealthy cases to 14 percent among the very wealthy, with the overall list rate around 34 percent. The weighting procedure effectively scales up the weights of responding households to compensate for similar households that did not respond.

What Information Is Collected

Interviews for the SCF are conducted in person. Because of the detailed nature of the survey questions, the interview takes an average of 90 minutes to complete, and can range up to three hours for households with complex finances. The interview is generally conducted with the financially most knowledgeable member of the household, although other household members may also provide information (for example, the respondent's spouse
may be asked questions about his or her employment). Respondents are encouraged to consult records, and 30 to 40 percent do.

The survey begins by collecting basic demographic information on all household members, including their age, sex and marital status. The respondent is then asked to list the financial institutions at which household members have accounts or loans, including the type of institution, the way of doing business with the institution and the distance between the institution and the home or workplace of the person who uses it most. As respondents describe particular accounts or loans during the course of the survey, this "institution roster" is used to identify the institution at which each item is held.

The survey then goes on to collect detailed information on the household's financial assets, nonfinancial assets and liabilities. The section on financial assets includes checking, saving, money market, and call accounts; certificates of deposit; IRA and Keogh accounts; stocks; bonds; mutual funds; savings bonds; cash value life insurance; and trusts, annuities and other managed assets. For each item the respondent mentions, he or she is asked about its value and the institution at which it is held. Nonfinancial assets include the household's principal residence, investment real estate, vehicles, business interests, and other valuable assets like art and precious metals. Liabilities specifically mentioned in the survey include mortgages, home equity loans and lines of credit, loans for investment real estate, vehicle loans, student loans, consumer installment loans, and debt on credit cards. For each loan, the respondent is asked about the balance outstanding and other aspects of the loan's terms, including its duration, the interest rate, the typical payment, and the institution.

In addition to the core questions on assets and liabilities, the survey also collects information on other topics relevant to understanding households' financial situations, including the employment and pension coverage of the respondent and spouse; household income and tax filing status; coverage by health, life and disability insurance; the educational attainment of the respondent and spouse; the health status of the respondent and spouse; experience in applying for loans in the past five years; recent problems making payments on loans; and attitudinal data on risk, borrowing and saving.

Future pension benefits are often not known in detail by respondents, but represent a potentially important financial resource. To obtain more information on pension rights, the 1983 and 1989 SCFs included supplemental "Surveys of Pension Providers," in which employers of households reporting pension coverage were contacted and asked for their pension plan's Summary Plan Description. The Summary Plan Description is a legal document describing the features of the pension plan, such as contribution information and calculation of benefits. The information from these documents was coded, and made available as a separate file that can be merged with the household data.

Data Quality and Imputation

An important problem in wealth surveys concerns nonresponse to questions about values of assets and liabilities. Questions about ownership of specific assets or liabilities usually have response rates around 99 percent. Questions about the values of salient items--like the value of the principal residence or the balance in the main checking account--also have reasonably good response rates, exceeding 90 percent. But for other items, response rates on value questions range between 70 and 85 percent. The lowest response rates occur for stocks, cash-value life insurance, and business interests--items often held as long term investments, and whose values are not regularly checked.

The SCF takes a number of steps to address problems of missing or incomplete information. Beginning in 1995, the SCF started systematically asking respondents to indicate a range when they were unable to provide an item's exact value. In cases where respondents did not provide a value or range, the SCF imputes missing values using multiple imputation (see Kennickell, 1991). In brief, each variable is imputed multiple times, using an iterative, multivariate procedure. Compared to other imputation methods, multiple imputation has the advantage of preserving the second moments of the data, and providing a basis for measuring variance due to imputation.

Several studies have compared the SCF with other sources of survey data on wealth (for detailed comparisons, see Avery, Elliehausen and Kennickell, 1988; and Curtin, Juster and Morgan, 1989). According to these studies, the estimates of median wealth are similar across surveys, suggesting that the wealth holdings of the typical household can be characterized from shorter surveys without special sampling. However, there are two ways in which the SCF is quite different from other surveys. First, largely due to the improved representation of wealthy households, the estimate of mean wealth from the SCF is considerably higher than estimates from other surveys. By extension, the estimate of total wealth held by households is also much higher; indeed, total values estimated from the SCF generally compare favorably with aggregate data sources like the Flow of Funds, especially when differences in definitions are taken into account (see Antoniewicz, 1996). Second, compared to other survey
sources, the SCF shows a somewhat different picture of the distribution of wealth. As would be expected, the SCF shows a larger share of households with high levels of wealth, and higher holdings for such households. But it also shows fewer households with low or no holdings, probably because the more detailed questioning elicits more reporting.

Research Uses of the SCF

Household Saving

The SCF has played an important role in the analysis of household saving. From the SCF and other survey data, Bosworth, Burtless and Sabelhaus (1991) established that the 1980s decline in U.S. personal saving did not result from demographic shifts, but rather occurred for all demographic groups. Sabelhaus and Manchester (1995) compared the wealth of baby boomers with the wealth of their parents' generation at a similar age, finding greater wealth among boomers. Bernheim and Scholz (1993) compared actual age-wealth profiles with simulated profiles that would provide adequate saving for retirement. Gale and Scholz (1994) examined the effects of IRAs on household saving, finding some portfolio reshuffling but little increase in saving. Starr-McCluer (1996) and Carroll, Dynan and Krane (1998) used SCF data to examine precautionary motives for saving.

The SCF has also figured prominently in research on the effects of public and private pensions on saving. Using the 1962-63 data, Feldstein and Pellechio (1979) found a negative relationship between Social Security wealth and household savings. Subsequent studies have reexamined this issue, also covering the effects of pension wealth (see Gale 1998 for a recent contribution). Mitchell (1988) used the 1983 Pension Provider data to investigate worker knowledge of pension benefits, finding significant discrepancies between workers' and employers' descriptions of pension provisions. Starr-McCluer and Sundén (1998) have somewhat different findings, using the 1989 data.

Household Portfolios


Borrowing and Liquidity Constraints

The SCF has also been used to study borrowing and liquidity constraints. Whereas most studies have measured liquidity constraints indirectly, Jappelli (1990) and Cox and Japelli (1993) analyzed information provided by households on their experience with credit rejections. Gropp, Scholz and White (1997) studied the effects of bankruptcy laws on households' use of credit. The SCF has also been used to try to explain why households borrow on credit cards, despite the high interest rates (Calem and Mester 1995).

Wealth Inequality

Finally, the SCF has been a main source of information on changes in the distribution of wealth. Estimates of wealth concentration depend importantly on sample weights: Whereas several earlier studies found an increase in wealth inequality in the 1980s, more recent analysis using updated weights suggests little change in wealth inequality from 1983 to 1992, but a significant increase in 1995 (Kennickell and Woodburn 1997). Other studies have examined the role of pensions and Social Security in the distribution of wealth (Feldstein 1976b, Kennickell and Sundén 1997).

Getting the SCF Data

Data and documentation for the SCF may be downloaded from the Federal Reserve Board's website. The SCF data are also available for member institutions through the Inter-University Consortium for Political and Social Research (313-764-2570) and for the general public through the National Technical Information Service (703-487-4763). The Pension Provider data can also be obtained through the Board's web site, ICPSR, or the NTIS. For
questions about the survey, contact Gerhard Fries at the Federal Reserve Board, (202) 452-2578, or via e-mail at gfries@frb.gov.

Acknowledgements

We are grateful to Arthur Kennickell, Spencer Krane, Myron Kwast, Dean Maki, Cathy Montalto, Kevin Moore, and Amy Stubbendick for valuable comments, and to Jeanne Hogarth for her exceptional initiative and help in putting together the SCF session. The views expressed in this paper are those of the authors, and not necessarily those of the Federal Reserve Board or its staff.

References


Endnotes

1. Statistician, Division of Research and Statistics.
2. Senior Economist, Division of Research and Statistics.
3. Economist, Division of Research and Statistics.
4. Several other federal agencies have provided valuable support for the survey. The Statistics of Income Division of the Internal Revenue Service has played an essential role in the sample design. The survey was conducted by the U.S. Census Bureau in 1962 and 1963, the Survey Research Center at the University of Michigan from 1983 to 1989, and the National Opinion Research Center at the University of Chicago since 1992.

5. Other surveys of the same name were conducted from the late 1940s through 1971, and again in 1977, but were primarily concerned with gauging consumers’ buying intentions.

6. Other surveys containing information on wealth include the Survey of Income and Program Participation, the Panel Survey of Income Dynamics (1984, 1989 and 1994), and the Health and Retirement Survey.

7. See Kennickell, McManus, and Woodburn (1996) for additional discussion of the sample design.

8. At a given dwelling, the main family—or “primary economic unit” (PEU)—is the economically dominant single individual or pair of individuals (married or living as partners) and all other persons who are financially dependent on that person or couple. Almost all of the financial information collected in the survey refers to the PEU, although some questions are asked about the finances of secondary household members.

9. The file is made available by the Statistics of Income Division of the IRS under strict rules governing confidentiality, the rights of potential respondents to refuse participation in the survey, and the types of information that can be made publicly available.

10. Kennickell and McManus (1993) discuss the correlation between the wealth proxy and net worth as measured in the survey.

11. Most statistical software packages have options for computing summary statistics on a weighted basis. The sum of the sample weights equals the number of households in the U.S.: 93.1 million in 1989, 95.9 million in 1992, and 99.0 million in 1995.

12. Note that nonresponse reflects not only refusals, but also inability to get around the security and other barriers protecting the privacy of wealthy families. Before list respondents are contacted by interviewers, they receive a letter informing them that they have been selected as a respondent, describing the purpose of the survey, and explaining that the interview is voluntary and confidential. Interviewers also convey this information when they contact respondents in person.

13. Some phone interviews are conducted at the convenience of the respondent. Starting with the 1995 SCF, interviewers have used laptop computers to administer the questionnaire.

14. In the data files, data collected on the respondent and spouse have been mapped into variables for the "head" and "spouse," where the head may not be the same person as the respondent. In PEUs based on a couple, the "head" is taken to be the male, or the older individual in a same-sex couple. In other PEUs, the respondent is the head. The term "head" is purely a convention related to the organization of the data and implies no judgment about the structure of families.

15. For analysis of the institution data, see Ellliehausen and Wolken (1992) or Kwast, Starr-McCluer, and Wolken (1997).

16. For analysis of the data on debt holdings, see Canner, Kennickell and Luckett (1995).

17. While respondents have long been able to select a range from a "range card," the use of computer-assisted interviewing permits more extensive range information. See Juster and Smith (1997) and Kennickell (1997) on the advantages of range data.

18. Multiple imputation has been used since the 1989 SCF. Each variable in the data file has a "shadow" variable that describes the original state of the variable, including whether it was missing for any reason. In 1983 and 1986, values were imputed using a variety of methods, including regression and hot deck.

19. Each variable is imputed as a function of variables with which it is likely to be correlated. Since 1989, the SCF data files have consisted of original records repeated five times, with each of these "implicates" imputed independently. To compute descriptive statistics, it is best to use all five implicates, dividing the
sample weights by five. Standard errors due to sampling can be computed using replicate weights (provided in a separate file), and standard errors due to imputation can be calculated from the five individual estimates from the implicants (see Kennickell, Woodburn, and McManus 1996; the SCF codebook provides details). For econometric analysis, it is common to use one implicate only, although in principle the standard errors should be adjusted for imputation to avoid biased significance tests. See Montalto and Sung (1996) for a useful discussion.

20. It is also important to note that the SCF's special sample design lowers the statistical variability of estimates of mean or total net worth.


22. The Board's website is http://www.bog.frb.fed.us. To get to the SCF home page, select the option "Surveys" under "Domestic and International Research."