

The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances

Arthur B. Kennickell
Federal Reserve Board
Mail Stop 153
Washington, DC 20551 USA
Arthur.Kennickell@frb.gov

1. Design of the Survey of Consumer Finances

The Survey of Consumer Finances (SCF) is a triennial cross-sectional survey of U.S. households, sponsored by the Federal Reserve Board, in cooperation with the Statistics of Income Division (SOI) of the Internal Revenue Service. Since 1992, data for the survey have been collected by NORC, a social science and survey research organization at the University of Chicago. The survey began in 1983 and had a substantial revision of both the questionnaire and sample design in 1989. Since that time, the survey has changed marginally and usually in reaction to developments in the financial market place or in response to measurement problems.

The SCF is designed primarily to collect information on the assets and liabilities of U.S. households, their use of financial services, their employment history and pension rights, as well as their demographic characteristics, attitudes and other characteristics.¹ The survey is widely used to study a range of issues for which wealth and financial data are important. In some cases, such research focuses on the behavior or experience of individuals, but in other cases the focus is more on overall market outcomes. Because of the range of purposes for which the SCF data are used, it is particularly important that the survey adequately represent the full distribution of wealth in the U.S. Because the wealthiest 1 percent of households is estimated to hold about a third of all household net worth, it is critical that the SCF pay particular attention to that rarified group.² At the same time, the survey is expected to provide adequate representation of much less wealthy households as well.

The sample design for the SCF is the most basic foundation for the data collection (see Kennickell (2005a)). The survey employs a dual-frame design, including an area-probability and a list component. The area-probability (AP) sample is selected from a geographically based national frame developed by NORC at the University of Chicago (O'Muircheartaigh *et al.* (2002)). The AP sample is selected in three stages. At the first stage, areas of the country (metropolitan areas and rural counties) are selected using a stratification scheme to balance the sample along key dimensions. At the second stage, sub-areas are chosen, again using stratification to balance the sample. Households are chosen using systematic sampling from address listings within the sub-areas such that every household in the overall sample is selected with equal probability. The AP sample provides robust coverage of the nation and good representation of behaviors that are broadly distributed in the population. About two-thirds of the ultimately completed cases derive from this sample.

The list sample is used to over-sample households that are likely to be relatively wealthy. The basis of the sample is a set of specially edited individual income tax returns developed by SOI, primarily for use in modeling in support of tax policy by the Office of Tax Analysis at the U.S. Treasury and the Joint Committee on Taxation of the U.S. Congress, as well as various other researchers. The individual records contain much of the information that would be present on the forms submitted as a part of the annual reporting of income for tax purposes. The SCF selects observations in two steps. In the first stage, observations in areas selected for the first stage of the AP sample are selected; this is a practical accommodation intended to control costs on the survey.³ The remaining cases are stratified using a model of wealth conditional on the variables in the SOI data that blends two approaches. One part models wealth as a simple grossing up of capital income flows. The other part is actually estimated using an anonymized match

¹ For a review of some highlights of the 2004 SCF and references to technical documentation, see Bucks, Kennickell and Moore (2006).

² See Kennickell (2006) for an extended discussion of the wealth distribution in the U.S.

³ Wealthy households appear not to be distributed the same way as population in general (Frankel and Kennickell (1995)). Although a substantial number of unusually wealthy people live in relatively thinly populated areas, they are more concentrated in the largest metropolitan areas than people in general. Nonetheless, the available evidence suggests that the first-stage selection of list sample cases does not induce serious distortions in the survey.

of sample and survey data for the previous round of the survey. An important detail is that three years of data are used for each SOI case in order to smooth out transitory variations in income that might otherwise distort the estimation of the model. Similarly, the most current three years of data are used in computing both parts of the wealth model. A “wealth index” is calculated for all tax filers by blending the two estimates, and this index is used to classify the cases in the sampled areas into seven strata, which are defined using percentiles of the distributions of the index to ensure consistency of the stratum boundaries over time.⁴ Cases within the areas selected for the AP sample are selected using a systematic sampling approach within each stratum. Higher strata are sampled at progressively higher rates.⁵

Only the larger geographic overlap is common between the AP and list samples. Otherwise, selection is entirely independent. Although this independence raises some complications at the weighting and analysis stages, it provides a useful means of examining the importance of the list sample to the survey. The remainder of the paper focuses on a few key contributions of the list sample.

2. The Contribution of the List Sample

The list sample serves two main functions. First, it provides a basis for more precise estimates of wealth in general and of narrowly held assets than would be possible in a less-structured sample without a much larger sample size. Second, the structure of the over-sample provides a means of correcting for nonresponse, which is differentially higher among the wealthy; thus it provides a means of correcting for nonresponse bias in wealth estimates.

About 98 percent of SCF cases with at least \$5 million of net worth in 2004 derived from the list sample; more than 85 percent of cases with at least \$1 million of net worth and about 75 percent of the cases with at least \$500 thousand dollars came from this sample. Thus, it is clear in an informal sense that the list sample adds substantially to analysis that depends on good representation of the upper tail of the wealth distribution. Such a gain might, in principle, be attained if the AP sample were sufficiently large or if there were a reliable way of screening households for their wealth. The efficiency gain will be quantified further.

Virtually all non-mandatory surveys suffer from nonresponse. If the distribution of outcomes for nonrespondents is the same as the distribution for respondents, then nonresponse has only the effect of reducing the sample size available for analysis—similar to the effect of random subsampling of the original sample. In some cases this may be true, but it is at the least a questionable assumption in the absence of evidence. In a wealth survey, the sensitivity of the subject and time cost for people with complex assets to be interviewed should be enough to raise *a priori* concerns. For the SCF, there is evidence (Kennickell (2005b)) of complex systematic effects. For example, response rates decline with capital income and rise with age and amounts of charitable contributions made. In the stratum of the SCF list sample that contains the respondents likely to be the wealthiest, the overall response rate is only 10 percent. The survey has often been criticized for this low cooperation rate. Regrettable as this rate is, the fact that it is known is actually a strength of the survey. Presumably, other surveys also have a similar problem, but without some means of identifying it, they will fail to correct for an important source of bias in the estimation of wealth.⁶ In the SCF the original frame data for the list sample provides a rich basis to use for adjusting the sampling weights to compensate for nonresponse.

⁴ The sample specifically excludes people who are listed as being members of the *Forbes* list of the 400 wealthiest people in the U.S. According to Kennickell (2006), if the data for this group are taken at face value, they account for about 2 percent of household net worth. However, the *Forbes* data may have measurement problems that complicate their use. Some of the figures they report are amounts owned by extended families or by charitable trusts controlled by (but not legally owned by) a person. SCF data suggests that they also miss some people who are as wealthy as other people included in the list. The argument for excluding this group from the SCF sample is that because such people are typically surrounded with levels of staff intended to keep other people away, they would be extraordinarily expensive to attempt to interview, and the success rate could reasonably be expected to be quite low. Moreover, because these people are so well known, it would be almost impossible to protect their confidentiality without destroying the statistical utility of the data they would provide.

⁵ Neither the stratum boundaries nor the sampling rates can be revealed. However, it can be said that the highest three strata correspond approximately to about the wealthiest two percent of tax filers.

⁶ Because as a part of its interviewer training and field procedures the SCF devotes substantial time and money toward dealing with cooperation problems among the wealthy, the success rate for other less specialized approaches would be likely to be lower than in the SCF.

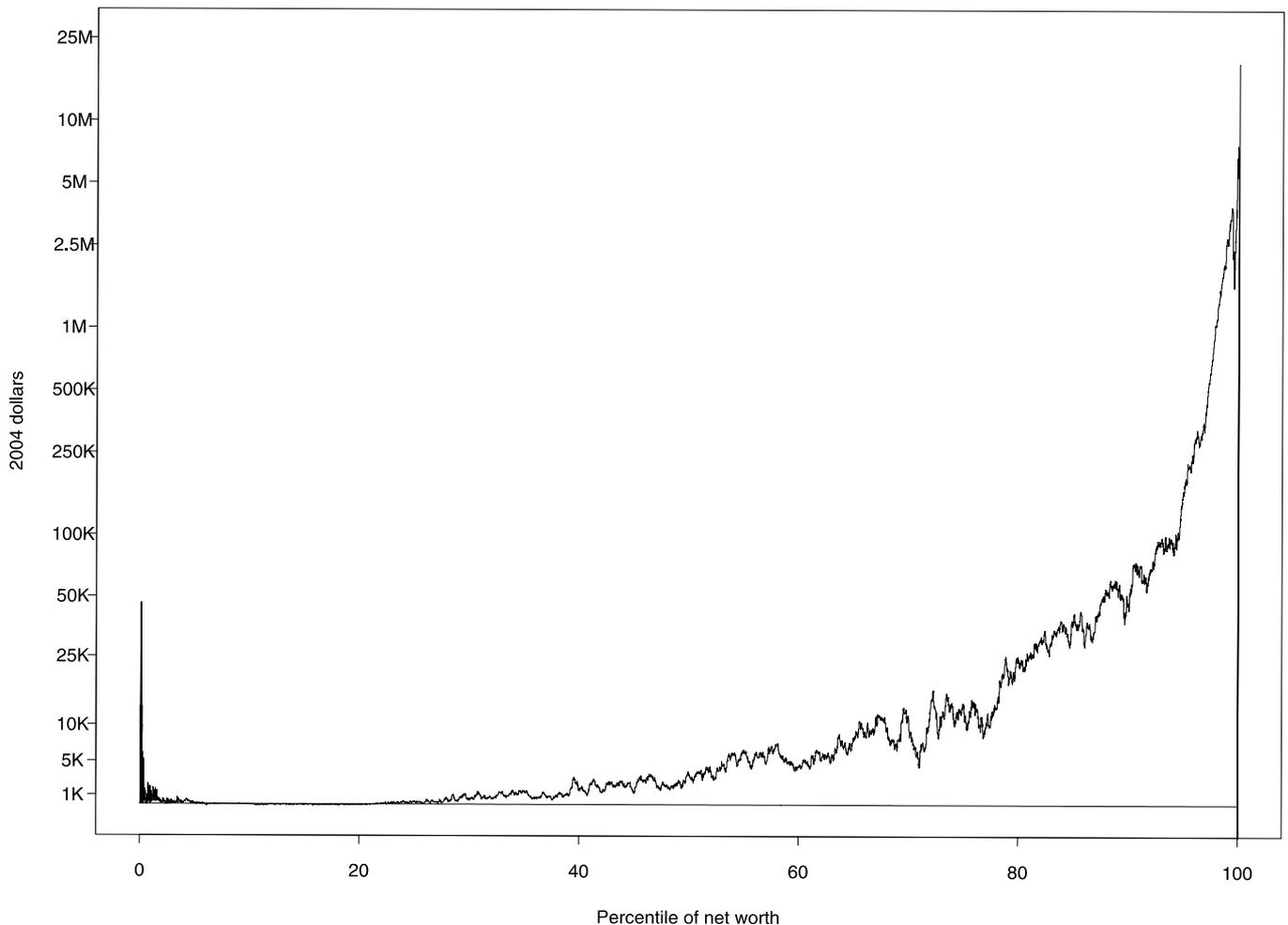


Figure 1: Distribution of net worth in 2004 based on full sample minus distribution based on area-probability sample only, 2004 dollars, by percentile of net worth.

A good sense of the importance of the bias correction is given by the information in figure 1, which shows the difference at each point in the distribution of net worth between the distribution estimated using only the AP sample with the weights for the AP sample adjusted using all factors available for that sample and the distribution estimated using the full sample with all the full set of nonresponse adjustments. For very low values, the addition of the list sample appears to make the distribution become more negative. However, closer examination of the values in this range reveals that the plot is merely reflection noise in that region; there is no systematic effect there. From about the 20th percentile and higher, the level of net worth at each percentile under the combined AP and list samples is higher than under the AP sample alone. The approximate increase in the level of the wealth distribution from integrating the list sample is \$250 (1.9%) at the 25th percentile, \$3,500 (3.9%) at the median, \$13,600 (4.3%) at the 75th percentile, \$43,600 (5.5%) at the 90th percentile and \$2,661,000 (74.0%) at the 99th percentile. At the top of the distribution, the list sample fills a niche that very thinly populated with AP sample cases. By “displacing” the top of the estimated AP distribution downward, the inclusion of the list sample would tend to make the values associated with lower percentile points higher. If one “synthetic case” with a weight equal to one percent of the population and a net worth equal to the 99th percentile value under combined samples, this is sufficient to approximately reproduce this pattern. However, the list sample does not simply displace the top one percent of the distribution; it also affects the distribution by recalibrating the percentiles of the distribution by adjusting for under-representation of households in regions below the very top.

The second benefit of the list sample is that it makes possible more stable estimates of quantities that are relatively strongly affected by the upper tail of the wealth distribution. If the SCF were only concerned with such estimates, it would attempt to sample each dollar of wealth (or some other such variable) with equal probability. But the survey must serve a number of different purposes, including addressing issues largely relevant to the lower part of the wealth distribution. Thus, the definitions of the list sample strata differ from what would emerge from an optimal stratification calculation based on total wealth. Table 1 shows the estimate of the shares of total net worth held by different groups defined by percentile ranges of

the distribution of net worth, along with standard errors for those figures. The estimates are presented for the full sample and for the AP sample alone. The standard errors for the AP sample estimates have been rescaled downward to account for the overall difference in the sample sizes of the two samples. Despite the large difference in the levels of wealth at the top of the wealth distribution under the two samples, the *shares* of the wealthiest 1 percent are very similar—about one-third of the total. But the standard errors are quite different—5.1 under the AP sample and 1.2 under the combined samples. The differences in standard errors are less pronounced for other groups, but the standard error under the combined sample is uniformly smaller.

The list sample also makes it possible analysis of portfolio patterns that could not only be supported by an enormous AP sample, assuming the absence of nonresponse bias. For example, in the 2004 SCF, only 1.8 percent of households had direct holdings of government bonds (other than savings bonds) or commercial bonds. Out of the approximately 400 cases with bonds in the combined samples, only about 10 percent of the cases were from the AP sample.

Table 1: Net worth shares in 2004 for various percentiles of the net worth distribution; area-probability sample only and combined area-probability and list samples; percent.

<i>Net worth group</i>	<i>AP sample only</i>			<i>AP and list samples combined</i>		
	<i>Share of group</i>	<i>SE of share*</i>	<i>Number of observations</i>	<i>Share of group</i>	<i>SE of share</i>	<i>Number of observations</i>
Lowest 50%	2.9	0.3	1,642	2.5	0.1	1,741
50%—90%	30.9	2.8	1,097	27.9	0.9	1,343
90%—95%	12.9	1.1	132	12.0	0.7	269
95%—99%	19.7	1.4	109	24.1	1.2	454
Highest 1%	33.6	5.1	27	33.4	1.2	715
All	100.0	0.0	3,007	100.0	0.0	4,522

* Standard errors for the area-probability sample estimates are reduced by $\sqrt{3007}/\sqrt{4522}$ (0.817); see text for discussion.

REFERENCES

- Bucks, Brian K., Arthur B. Kennickell and Kevin B. Moore [2006] “Recent Changes in Family Finances: Evidence from the 2001 and 2004 Surveys of Consumer Finances,” *Federal Reserve Bulletin*, pp. A1-A38.
- Frankel, Martin and Arthur B. Kennickell [1995] “Toward the Development of an Optimal Stratification Paradigm for the Survey of Consumer Finances,” <http://www.federalreserve.gov/pubs/oss/oss2/papers/frankel.kennickell.asa95.pdf>.
- Kennickell, Arthur B. [2006] “Currents and Countercurrents: Changes in the Distribution of Wealth, 1989-2004,” <http://www.federalreserve.gov/pubs/oss/oss2/papers/concentration.2004.5.pdf>.
- Kennickell, Arthur B. [2005a] “The Good Shepherd: Sample Design and Control for Wealth Measurement in the Survey of Consumer Finances,” paper presented at the January 2005 meeting of the Luxembourg Wealth Study, <http://www.federalreserve.gov/pubs/oss/oss2/papers/sampling.perugia05.2.pdf>.
- Kennickell, Arthur B. [2005b] “Darkness Made Visible: Field Management and Nonresponse in the 2004 SCF,” <http://www.federalreserve.gov/pubs/oss/oss2/papers/asa2005.5.pdf>.
- O’Muircheartaigh, Colm, Stepahie Eckman, Charlene Weiss [2002] “Traditional and Enhanced Field Listing for Probability Sampling,” *Proceedings of the American Statistical Association Social Statistics Section*, pp. 2563-7.

RESUME

The Survey of Consumer Finances (SCF) is intended to be a survey of the assets and liabilities of U.S. families, but there is a particular need to come as close as possible to representing well the full distribution of wealth. To this end, the survey employs a dual-frame sample, one part of which is based on area-probability sampling and the other uses data from administrative data to over-sample people likely to be wealthy. The over-sampling serves two main functions. First, it provides more precise estimates of wealth in general and of narrowly held assets than would be possible with a less-structured sample of larger size. Second, the structure of the over-sample provides a means of correcting for nonresponse, which is differentially higher among the wealthy. This paper provides a brief discussion of the SCF sample and it provides examples of the gains from over-sampling by comparing estimates and standard errors of estimates derived using the full sample with those obtained by using the area-probability sample alone.