Estimation of Household Net Worth Using Model-Based and Design-Based Weights: Evidence from the 1989 Survey of Consumer Finances

Arthur B. Kennickell
Senior Economist and Project Director Survey of Consumer Finances
Board of Governors of the Federal Reserve System
Mail Stop 180, Washington, DC 20551

R. Louise Woodburn
Mathematical Statistician
Statistics of Income Division, Internal Revenue Service
P.O. Box 2608, Washington, DC 20013-2608

April 1992

Opinions expressed in this paper are the responsibility of the authors alone and in no way reflect the views of the Board of Governors, the Federal Reserve System, or the Internal Revenue Service. The authors gratefully acknowledge the methodological comments and suggestions of Roderick Little, Donald Rubin, and Fritz Scheuren. The authors are also grateful for other comments from B.K. Atrostic, Edward Ettin, Myron Kwast, and Janice Shack-Marquez. The authors bear all responsibility for any errors remaining in this paper.

I. Introduction

In this paper, we present estimates of the distribution of net worth across U.S. families based on information from the 1989 Survey of Consumer Finances (SCF) as an illustration of the use of design-based and model-based weights. For comparison, we produce similar estimates from the 1983 SCF. In order to assess the precision of our estimates of the 1989 net worth distribution, we present calculations of standard errors due to sampling in the case of the model-based weights. At this point, we are not able to compute standard errors for the design-based weights. The material presented here represents a progress report of on-going work on weighting.

We provide a decomposition of the 1989 net worth distribution by percentiles of the distribution and by composition of assets and liabilities held by different groups. Using the 1989 SCF data and a set of model-based weights we have developed, we estimate the aggregate total net worth to be 15.1 trillion 1989 dollars with an estimated standard error due to sampling of 0.70 trillion 1989 dollars. The wealthiest half percent of the population is estimated to hold 28.8 percent of the net worth with an estimated standard error due to sampling of 2.0 percentage points. Under the design-based weights, we estimate aggregate household net worth to be 17.0 trillion 1989 dollars and the proportion of net worth held by the top one-half percent to be 29.1 percent. For contrast, we use the 1983 SCF to estimate comparable figures. According to our estimates, total net worth represented by the survey population in 1983 was 10.2 trillion 1983 dollars and the top one-half percent of the net worth distribution held 24.1 percent of all household net worth. As we note in greater detail below, because we are not yet able to make standard error calculations for estimates computed using the design-based weights or for the

1983 figures, we cannot offer a formal statistical test of the significance of the change in the share held by the top one-half percent, though we do examine a range of possible tests.

The plan of this paper is as follows: in Section II, we provide a detailed technical description of the survey design and an overview of the data collected for the 1989 SCF. The next section discusses sample weight construction with emphasis on the development of replicate weights for variance computations. Section IV presents estimates of the distribution of household net worth along with the standard errors. Finally, in Section V, we summarize our findings and point toward further research.

II. 1989 Survey of Consumer Finances

The purpose of the 1989 SCF is to provide a broad, yet detailed, view of the financial position of households.[1] Comprehensive information was gathered on the assets and liabilities of each household interviewed. In addition, data were collected on current jobs and the current (and future) pension rights of respondents and their spouses, as well as such demographic characteristics as age and family composition. Care was taken at every stage of data generation to account for the particular problems involved in the collection of wealth data. In particular, interviewers were given special training on household finances and detailed editing specifications were developed.

Data for the survey were obtained between the months of July 1989 and March 1990 by the Survey Research Center at the University of Michigan. Thus, the survey might be thought of

[1] In this paper, the terms "family" and "household" are used interchangeably to indicate the "primary economic unit" within a traditional household. The primary economic unit is loosely defined as the economically-dominant person or couple within a household, their minor children, and any other people in the household who are economically dependent on that person or couple. A more detailed description of the survey is given in Kennickell [1991a and 1991b] and Kennickell and Shack-Marquez [1992].

as offering a picture of family finances at the end of 1989. Interviews were mainly conducted in person and averaged about 75 minutes. Item nonresponse rates are roughly comparable to those found in other economic surveys. The missing information was imputed statistically using techniques of multiple imputation and Gibbs sampling.[2]

The sample design of the 1989 SCF is complex.[3] The sample falls into two major parts: an overlapping panel cross-section based on the 1983 SCF design and a new cross-section sample. The overlapping panel cross-section is most easily thought of as two samples, a panel of individuals and a panel of addresses. A sample of individual respondents was followed from the 1983 SCF. In addition, the original address sample for the 1983 SCF was augmented to account for new construction, and a systematic selection of those addresses was approached for interviews. In cases where an original 1983 respondent had not moved, the interview of the respondent has both longitudinal and cross-section representation.[4] In this paper, we use only the cases that have 1989 cross-section representation.

The new cross-section sample is a dual-frame design with an area probability component and a list component. The area probability frame is based on geographic information and the sample from that frame was drawn using standard multi-stage area-probability sampling procedures. The list frame consists of a sample of administrative records maintained by the

[2] A more thorough discussion of the problems of missing data in the survey and a description of the imputation procedure is given in Kennickell [1991a].

[3] A more complete description of the sample design is available in Heeringa and Woodburn [1992].

[4] The sample of individuals also includes a supplementary sample of administrative records for the 1983 SCF. Because the cross-section representation of these families is unknown, these families are taken to have only panel representation.

Statistics of Income Division (SOI) of the Internal Revenue Service. These records contain data coded from individual income tax returns filed in 1988 (generally for tax year 1987).[5]

The list frame was stratified by a "wealth index" designed as a proxy for net worth and units were selected disproportionately in order to oversample units with higher values of the wealth index. The wealth index was constructed for each list frame element using reported income flows capitalized at various rates of return and other information.[6] Work reported in Kennickell and Woodburn [1992] suggests that the index is a "noisy" representation of net worth, though it appears to be a better proxy for gross assets.

The achieved cross-section sample from all parts of the design includes 3,143 families, of which 866 come from the list frame. The area-probability cases were approached directly by interviewers and the response rate for these cases was about 69 percent. In contrast, the list-sample cases were given a prior opportunity to refuse participation by returning a postpaid card. About 36 percent of the original sample of list cases refused participation at this stage by returning the card. The remainder of the list cases were approached by interviewers, yielding an overall interview rate for the list sample of about 34 percent.[7]

---

[5] For a description of the Statistics of Income Individual Program in place when the SCF selections were made, see Internal Revenue Service [1990]. Statistical and research uses of SOI data are closely regulated to guarantee that individuals (and other entities) will remain protected against any disclosure of their financial and tax data (e.g., Wilson and Smith [1983]). For the 1989 SCF, contractual agreements between the Federal Reserve Board, the Survey Research Center, and SOI clearly specify the limitations on the use of the administrative data in order to guarantee the privacy rights of the individual taxpayers.

[6] For example, taxable interest income was assumed to be supported by a stock of interest-bearing assets equal to ten times the interest income (implicitly this assumes a rate of return of ten percent).

[7] Preliminary information on unit nonresponse for the list sample in the 1989 SCF is given in Woodburn [1991]. The results of her investigations are encouraging, suggesting that, despite the high nonresponse rates, the response bias is minimal for key variables. Additionally, she shows that weight adjustments that compensate for different rates of selection and response, as employed for the model-based weights described in this paper, are highly effective at eliminating selection and response bias.

While the interview rate for the list cases is not high according to usual criteria, this figure merely makes explicit what is hidden in other household surveys without such a sample or other auxiliary information to identify the problem. Moreover, as we note in detail below, we can make systematic nonresponse adjustments to the SCF sample by estimating response models using information from the list frame.

III. Development of Sampling Weights

Because the SCF is based on a complex stratified dual-frame sample design and because unit nonresponse was substantial, sampling weights are an essential element in computing estimates of the net worth distribution, or virtually any other distributions based on the survey data. To assess the variability of our calculations with respect to alternative assumptions about weighting, different methodologies were used to develop the weights, resulting in design-based weights and model-based weights. The design-based weights are a later generation of the weights used in the calculations reported in Kennickell and Shack-Marquez [1992]. The methodology used to develop the design-based weights is described in detail in Heeringa and Woodburn [1992]. The general methodology used for the weight development is given below. A more detailed description of the computation of the model-based weights is given in the Appendix.

In principle, one might expect the model-based weights to provide better point estimates since they incorporate more information than the design-based weights. For this reason we take the model-based weights as our main weights. To facilitate sampling variance calculation, bootstrap replicate weights were computed for these weights as well.

The methodologies used to develop the two types of weights share a common framework. Calculation of both types of weights proceeds in three main stages:

1. A probability of observation is computed for each case under both the area-probability and list designs.
2. The joint probability of observation under either frame is computed.
3. The resulting weights are subjected to post-stratification.

The design-based weight and the model-based weight differ mainly at the first stage. In the spirit of Hartley [1962], a weight that pools the two samples entails knowledge of the probability of observation of each case under both designs. The chief difference in the two methodologies is the determination of the probability of observation for list and area-probability sample cases under the list design. The design-based weight is developed based on the original sample design, whereas the model-based weight employs models to estimate probabilities of observation.

For both classes of weights, the computation of the joint probability of observation under either frame is given by

(1) $$\Pr(\text{case i observed}) = \pi^i_{AP} + \pi^i_{LIST} - \pi^i_{AP}\ \pi^i_{LIST}\ ,$$

where π^i_{AP} is the probability of observation of a case within the area-probability frame and π^i_{LIST} is the probability of observation of a case within the list frame. The resulting weight associated with a case is the inverse of this probability.

The weights are then subjected to post-stratification.[8] Reliable estimates of marginal frequencies are available for the population from the March 1989 Current Population Survey by

[8] A good overview of post-stratification applications is given in Little [1991].

region and Metropolitan Statistical Area (MSA) status within region, as well as by home ownership status within age groups of the family head. All weights were subjected to three iterations of a raking procedure that controls to the distributions of age by homeownership and then to region by MSA status.[9]

The estimates computed using the 1989 SCF presented here challenge traditional variance estimation methodology. Without many simplifying assumptions, traditional variance estimators would not be appropriate given the complexity of the sample design. Also, we need to estimate variances for nonlinear and nonfunctional statistics -- for example, the proportion of wealth held by the wealthiest half percent of the population. In order to compute variances for such estimates, it is necessary to use a replication technique, such as the bootstrap, balanced repeated replication, or the jackknife.[10] In order to facilitate the estimation of variances due to sampling and the uncertainty inherent in the computation of the model-based weights, we have chosen to compute a set of replicate weights based on the bootstrap technique for the model-based weights.[11] We drew eleven bootstrap samples for this purpose using all available cross-section interviews as a base.[12] The procedure we applied to draw each bootstrap sample treats the 1983

[9] While full cell counts are also available for the cross-product of these two margins, some cell counts within the survey are too small for efficient post-stratification. See Oh and Scheuren [1978] for a description of the use of raking ratio estimation.

[10] An overview of replication techniques for variance estimation and applications in complex surveys can be found in Skinner et. al. [1989].

[11] This construction is described more fully in Kennickell and Woodburn [1992]. The variances computed here do not take into account all of the uncertainty due to imputation of missing values. We plan to eventually incorporate this uncertainty into the variance computations and to compute bootstrap replicate weights for the design-based weights as well.

[12] The number of bootstrap samples generally recommended is far larger than the 11 (12 including the full sample weight) used here. Applications of the bootstrap technique to complex survey settings is fairly new. Typically, such applications of the bootstrap involve the selection of 50 to 500 bootstrap samples to compute

panel cross-section, 1989 area-probability, and 1989 list samples separately and attempts to mimic the major sources of variation for each sample.

The original 1983 and 1989 area-probability designs define self-representing and non-self-representing areas which are treated as certainty strata and non-certainty strata, respectively, at the first stage of sampling. Each design includes one PSU per stratum for non-self-representing areas. For purposes of variance calculation, these PSUs have been paired by Steven Heeringa at SRC, separately for the 1983 and 1989 designs. The pairing was done by the major census regions and by MSA status. The 1989 area-probability and list samples cover the same set of PSUs. For the 1983 sample, the areas are largely different, with the principal exception of the self-representing areas.

For the non-self-representing areas for both the 1983 and the 1989 samples, we drew two areas from each of the area pairs using simple random sampling with replacement (SRSWR). Within each bootstrap replicate, the same non-self-representing areas were selected for both the 1989 area-probability and list samples. For the self-representing areas in the case of the area-probability samples, we selected segments within each area by SRSWR. For list-sample cases in the self-representing areas, respondents were stratified by wealth index strata and by the pseudo-

variances and confidence intervals for a specific estimate. We have chosen to actually compute weights to correspond with our bootstrap samples and to provide these weights to the data users as a variance computation tool. For the preliminary variance computations presented in this paper, we use the 12 different weights to develop tests with approximately 11 degrees of freedom. The choice of computing only 11 bootstrap samples at this point was made to decrease the computational burden of the bootstrap weights. Also, the computational burden, both for us and the typical data user, of computing estimates using the bootstrap weights was taken into account.

replicate groups created at the final stages of the original sample selection, and unit selection was made by SRSWR within these strata.[13]

We repeated the procedures used for the construction of the model-based weights for each bootstrap sample to yield eleven replicate weights. Thus, for each bootstrap sample, we recomputed all of the necessary models and subjected each bootstrap sample separately to the post-stratification process. While we believe that the bootstrap weights provide an adequate representation of variance due to sampling, the theoretical properties of variance estimates based on bootstrap weights for complex surveys are still being developed.[14] The principal virtue of the bootstrap weights we have computed here is that they are straightforward to compute and they are relatively simple to use in estimation. In the next stage of our work, we plan to explore the computation of variances using Balanced Repeated Replication (BRR) or another such method with better-understood theoretical properties.[15]

IV. Model-Based and Design-Based Estimates of the Net Worth Distribution

The 1989 SCF provides the most recent representative survey information on the distribution of the net worth of all U.S. families.[16] Consistent survey measurements of the

---

[13] The first group is a certainty stratum, the second is a stratified equal probability sample from the six wealth index strata, and the remaining groups are samples from all but the lowest wealth index stratum.

[14] Rao and Wu [1988] compare the performance of the bootstrap with other resampling methods of computing variances in complex survey settings.

[15] Use of the BRR technique in complex survey settings is discussed in Rao and Wu [1988]. Implementing BRR for the 1989 Survey of Consumer Finances would lead to approximately 40 replicate weights. We note that there is only a small difference between the critical values that would be used for testing with 40 degrees of freedom as opposed to the tests conducted here.

[16] Net worth is defined here as the difference between total assets and total liabilities, where total assets includes properties, vehicles, businesses, pension-type accounts from which withdrawals could be made, all other types of financial assets, and other assets, and total liabilities includes outstanding balances on credit cards, mortgages, lines of credit, traditional consumer credit, and other types of loans.

concentration of household net worth are available only for selected years since 1963.[17] Avery, Elliehausen and Kennickell [1988] and Avery and Kennickell [1988] have made earlier estimates of concentration for 1963, 1983, and 1986 based on a comparable series of household surveys. It is difficult to make comparisons of such complicated statistics over time, particularly when the survey measurement process is still evolving. However, the information collected in each of these years is sufficiently comparable that we feel comfortable with comparisons of data items at the level of aggregation presented here. There are also differences in sample design and weight construction. Our belief is that while changes in design are important, the sample designs have become more efficient over time and weighting has become more sophisticated, yielding more precise estimates.

One significant limitation of our work is that currently we can only make estimates of standard errors due to sampling for the the statistics we report based on the 1989 SCF. Thus, the only year for which we have a formal measure of the precision of our estimates is 1989. Each 1989 point estimate presented here is computed as the mean of the 11 bootstrap estimates and the full sample estimate.[18]

---

[17] Information on the distribution of wealth is also available from estate tax data (see Johnson and Schwartz [1990]). While the estate data are a useful time series, it is difficult to compare concentration estimates from that source with survey data. First, estate data are collected for only the small part of the population required to file a federal estate tax return. While one may reasonably make inferences about the shape of the upper tail of the net worth distribution, it is difficult to make statements about the shape of the entire distribution without perhaps "heroic" assumptions about the distribution of the unobserved wealth. Secondly, the unit of observation is the individual in the estate data, while the SCF collects wealth data on a family basis. Finally, it is not clear that the arrangement of assets observed at the time of an individual's death is an accurate reflection of that person's resources when alive. People may rearrange their assets in anticipation of death and some assets may be held in trusts that would not be recorded in an estate tax return. The difference between the survey and estate measures of net worth will be examined in detail in a future paper (Johnson and Woodburn [1992]).

[18] We use the full sample as a bootstrap sample for both computational convenience and also to increase the number of bootstrap samples. In theory, the expected value of the mean of the bootstrap samples is the full sample estimate for linear estimates. In the case of the 1989 SCF, the full sample estimate differs only slightly from the

There are many ways of characterizing changes in the distribution of net worth. One commonly used measure, the Gini coefficient, is defined as one minus twice the area under the Lorenz curve, which is a plot of the cumulative percentile distribution of net worth against the corresponding percentiles of the population. If all wealth were held by one individual, the Gini coefficient would equal one. If all wealth were held equally, the Gini coefficient would equal zero. Using the 1983 SCF, we estimate a value of 0.777 for this coefficient in that year. The 1989 SCF is the first of this series of surveys for which it has been possible to examine a range of conceptually different weights. For 1989, the model-based weights indicate that the Gini coefficient has risen to 0.793, suggesting that there had been an overall increase in the concentration of net worth in the upper tail of the distribution.[19] The estimate using the design-based weight is 0.800. Given the great differences in the construction of the weights, it is comforting that they yield such similar results at this level.

We would like to test the hypothesis that the Gini coefficient for 1989 is less than or equal to the 1983 Gini coefficient against the alternative that the 1989 Gini coefficient is strictly larger than the 1983 value. That is, we would like to compute the significance of a t-test that compares the hypothesis $H_0 = \{GINI_{89} \leq GINI_{83}\}$ vs. $H_1 = \{GINI_{89} > GINI_{83}\}$.[20] A one-tailed test is an appropriate test of whether the figure increased. Using the 11 bootstrap weights plus the

---

mean of the bootstrap samples. However, use of the full-sample estimate may cause the estimated sampling variance to be understated.

[19] The standard error due to sampling estimated under the model-based weight is 0.011. The range of the raw bootstrap estimates was 0.772 to 0.820.

[20] This test assumes that the statistic of interest follows a normal distribution and that the variances of the two estimates are similar.

full sample weight, we can compute a one-tailed t-test of H_0 against the alternative.[21] To perform this test, we need estimates of the variance of the estimates for both 1983 and 1989.[22] While we are able to estimate the sampling variance for the 1989 estimate using the model-based weights, currently we are not able to estimate the variance of the 1983 estimate and, thus, we cannot offer a formal test of the statistical significance of the change in the point estimate of the Gini coefficient.[23] However, we can at least examine the outcomes of a range of possible tests.

The list sample determines most of the variation in the estimation of the upper tail of the net worth distribution. The 1989 list sample was almost twice the size of the 1983 list sample. The variance of sampling error is proportional to the square root of the sample size. Thus, a reasonable baseline assumption might be that the 1983 variance is equal to 1.41 (the square root of two) times the 1989 variance. Under this assumption, the change in the Gini coefficient is statistically significant at the five percent level with a computed t-statistic of 3.09.[24]

---

[21] Formally, the test statistic is given by $(X_{89}-X_{83})/(S\ (1/n_{89}+1/n_{83}\ )^{1/2})$, where $S^2 =((n_{83}-1)VAR83 + (n_{89}-1)VAR89)/(n_{83}+n_{89}-2)$. Here n_{89} is equal to 12, the number of replicates used in the estimation of X_{89} and VAR89. For purposes of the hypothetical values we examine for VAR83, we assume that n_{83} is equal to n_{89}.

[22] It is important to note that the calculations that follow ignore any variation arising from the data values themselves -- i.e., all data values are treated as true constants. In fact, errors can arise in several ways through the data: errors may occur because of misunderstandings on the part of the interviewer or the respondent, because of processing errors, and because some data values are imputed, rather than reported. All of these types of errors should make the true variance of our estimates larger than the figures we estimate.

[23] The 1983 weight design differs from both the model-based weight and the design-based weights. However, the 1983 weight probably more resembles the model-based weight developed for the 1989 SCF. See Avery, Elliehausen and Kennickell [1988] for a discussion of the derivation of the 1983 weight used here. While we would like to recompute the 1983 weight using a procedure comparable to that used in the construction of the 1989 model-based weight and the corresponding replicate weights, this exercise would entail a major commitment of research time; nevertheless, since it is very desirable to have comparable weights available for all years of the survey, we expect to compute such weights eventually.

[24] If estimates of both variances were made in exactly the same way, this test would have 2 degrees of freedom. Intuitively it seems that this is an artificial inflation of the degrees of freedom. However, in the tests we pose, it makes no difference whether we use the significance level for 11 or 22 degrees of freedom. The five percent critical value for a t-test with 22 degrees of freedom is 1.717, and the critical value for the test with 11 degrees of freedom is 1.796.

One might argue that the differences in the 1983 and 1989 SCFs go beyond the sizes of the list samples. The 1989 list sample is more efficiently designed (arguing for relatively larger 1983 variance), and the area-probability sample in 1989 is smaller that that in 1983 (arguing for relatively larger 1989 variance). In addition, evidence presented in Avery, Elliehausen, and Kennickell [1988] suggests that the point estimate of the Gini coefficient was virtually unchanged between 1963 and 1983, though sampling error or other error could be causing the earlier data to give a false appearance of stability.[25] However, informally at least, this stability may suggest a relatively smaller value for the 1983 variance.

Under any assumption that the 1983 variance is less than 1.41 times the 1989 variance, the rejection of the null hypothesis will continue to hold. Considering the possibility that the 1983 variance is twice the 1989 variance, a one-tailed t-test rejects the null hypothesis at the five percent level with a computed t-statistic of 2.77. To fail to reject the null hypothesis at the five percent level, we have to assume that the 1983 variance is about five times the 1989 variance. While this result may seem strong, there are reasons we note below that one might still question that conclusion on a number of grounds.

Tables 1 and 2 provide more detailed estimates of the shape of the net worth distribution in 1989 using the model-based weights and the design-based weights, respectively. Table 1 also contains estimates of the standard errors due to sampling computed using the bootstrap sample weights for all of the figures presented. For comparison, Table 3 presents a table similar to one given in Avery, Elliehausen and Kennickell [1988] based on data from the 1983 Survey of

---

[25] If one defines comparable samples using the 1983 and 1986 SCFs, the estimate did not change appreciably between 1983 and 1986. See Avery and Kennickell [1990] for details.

Consumer Finances.[26] Standard errors are not currently available for the data reported in tables 2 and 3.

Under all the weighting schemes and in both 1983 and 1989, all assets and liabilities are disproportionately held by the wealthiest groups, but some assets are especially concentrated, particularly stocks, bonds, trusts, and businesses. Similarly, all types of liabilities are relatively concentrated in these groups.

According to the model-based weight, the top one-half percent of the net worth distribution held 28.8 percent of total family net worth in 1989 (with an estimated standard error due to sampling of 2.0 percentage points). The design-based weights indicate that this group held 29.1 percent of net worth. In 1983, the top one-half percent of the net worth distribution held 24.3 percent of total net worth.[27] Thus, estimates made using both the model-based weight and the design-based weight suggest that there may have been an increase in the share of wealth held by this top group in 1989.[28] Point estimates in the tables suggest that the increase in concentration in the highest net worth group was offset by a decline in the share of net worth held by the bottom 90 percent of the wealth distribution from 33.3 percent to 31.8 percent, and a decline for the 90 to 98.9 percentile group from 35.3 percent to 31.2 percent. Thus, the largest

[26] Due to limitations in data collection in 1986, the 1983 SCF provides the closest basis for comparisons to the 1989 SCF. The figures in table 4 differ slightly from those given in Avery, Elliehausen and Kennickell [1988] due to the use here of a more inclusive measure of net worth.

[27] It is worth noting that the 1989 estimates produced by every one of the bootstrap samples (a range of 25.9 to 32.9 percent) are larger than the 1983 figure.

[28] The estimated shift in concentration also holds if one makes broader cuts at the top of the distribution. According to the model-based weights, in 1989, the top one percent held 37.0 percent of net worth (with standard error due to sampling of 2.43 percentage points), and the top ten percent held 68.2 percent (with standard error due to sampling of 1.75 percentage points). In 1983, the corresponding point estimates for the top one and ten percent groups were 31.5 percent and 66.6 percent, respectively.

part of the apparent shift toward the top one-half percent came from a decrease in the proportion of net worth held by the remainder of the top ten percent of the net worth distribution.[29]

We would would like to test the statistical significance of the change for the top net worth group. However, we face the same problem in constructing a significance test for these changes that we faced with the Gini coefficient. We can only compute estimates of variances for the 1989 data. Nevertheless, as before we can compute tests under a range of assumptions.

If we take the same baseline assumption as in the case of the Gini coefficient -- that the main difference in the variance of the 1983 and 1989 estimates stems from the relative sizes of the 1983 and 1989 list samples -- then we have that the variance of the 1983 figure due to sampling error is about 1.41 (square root of two) times the variance of the 1989 figure. Under this assumption, we can reject with 95 percent confidence the null hypothesis that the proportion of net worth held by the top one-half percent of the net worth distribution has remained the same or declined. The computed t-statistic for this test is 4.68.

However, for the reasons discussed in the case of the Gini coefficient, one could make reasonable arguments that the 1983 variance should be greatly different based on the size of the total samples and the relative efficiencies of the designs in 1983 and 1989. In addition, as in the case of the time series of the Gini coefficient reported above, the earlier point estimates of the proportion of net worth held by the top one-half percent had been nearly constant since 1963,

[29] It is important to note that because we are not using panel data to make these calculations, the changes discussed should not necessarily be understood as applying to a fixed set of people. For example, when we talk about an increase in the concentration of net worth at the top of the net worth distribution, it could be that a fixed group became wealthier, it could be that the people who are estimated to be at the top of the net worth distribution in 1989 were farther down the distribution in 1983, or some combination could hold.

when this group was estimated to hold 24.6 percent of net worth.[30] As before, this apparent stability could be used to argue for a relatively smaller variance for the 1983 figure.

Any assumption of a 1983 variance smaller than our baseline assumption will also reject the null hypothesis. Alternatively, if we assume as before that the 1983 variance is twice the 1989 variance, the null hypothesis is still rejected at the five percent level of significance with a computed t-statistic of 4.20. To fail to reject the null hypothesis at the five percent level, we need to assume that the 1983 variance is more than 12.5 times the 1989 variance.[31]

However, one still might question these results arguing that insufficient allowance has been made for variation due to imputation of missing data or other problems in the reported data values (we have only attempted to deal with sampling error so far), that the sample designs are so different that comparisons are not valid, that bootstrap estimates of sampling variance are not good reflections of true sampling variance for the types of statistics we are estimating, that other tests might be more appropriate, or that other factors complicate or invalidate the comparisons. In addition, the comparison of 1983 and 1989 is complicated by the difference in economic conditions at each point: in 1983 the economy was at the beginning of an expansion, while in 1989 the economy was near the end of a long expansion. We intend to explore these questions in future papers. At this point, we have no additional scientific guidance to give the reader in forming an opinion about the significance of the estimated change.

---

[30] While the 1986 SCF is not directly comparable to the 1983 or 1989 SCFs, evidence presented by Avery and Kennickell [1990] indicates that the proportion of net worth held by the wealthiest group changed only very slightly.

[31] Preliminary estimates using the estate tax data support this finding. Using the wealth estimates derived from the estate data for 1982 and 1989, an increase is found for the proportion of the total wealth held by millionaires (as measured in constant dollars). More specific estimates using the estate data and comparisons with the SCF findings will be developed in Johnson and Woodburn [1992].

While the model-based weights and the design-based weights give similar estimates of the change in the concentration of net worth, they give different results for many of the detailed items shown in the table. Most importantly, the estimate of total net worth using the design-based weights is 12 percent higher than the estimate computed using the model-based weights. While most of the aggregate figures are lower under the model-based weight, there are two exceptions, bonds and trusts. This difference appears to be accounted for by the estimated greater holdings of these assets by the wealthiest group under the model-based weight.

An important question is how one should evaluate the differences in these results produced by the model-based and the design-based weights. In traditional sampling theory, sampling weights are not random variables and, thus, there is no theoretical basis for comparing results using different weights within that framework. That is, properties such as bias and variance are not associated with the weights themselves, but with specific estimates generated using the weights. Traditionally, one weight is deemed superior to an alternative weight based solely on the efficiency and unbiasedness of the resulting estimates.[32] Typically, estimates computed using design-based weights are deemed unbiased. However, the stability of estimates computed using design-based weights is weaker where nonresponse is significant and the sample design complex, as in the case of the 1989 SCF. As discussed in the appendix, a Bayesian framework provides a basis for the creation of the model-based weights. However, a methodology for testing which weight is "better," has not been developed. Nevertheless, there are two facts we believe are worth observing that provide at least informal support for the superiority of the model-based weights. The model-based weights incorporate more information

[32] For example, Holt and Smith [1979] provide a useful discussion.

about the representation in the population of the observed sample than do the design-based weights. In addition, we know from comparisons referred to in the appendix that the model-based weights do a better job of reproducing known population distributions of several income variables for the list sample than do the design-based weights.

V. Summary and Conclusion

As an example of the use of model-based and design-based weights, in this paper we have presented estimates of the distribution of net worth for U.S. families in 1989. The underlying estimation procedures are quite involved. Much of the paper has been devoted to providing a statistical foundation for the estimates we have presented.

Like earlier versions of the survey, the 1989 SCF is based on a complex sample design and unit nonresponse is substantial. For these reasons, weighting issues are very important. For the first time in the history of the SCF, we have had access to sufficient information to construct model-based weights incorporating a large amount of auxiliary information available for the list sample frame. We have also reported point estimates based on a more traditional design-based methodology.

Using both types of weights with the 1989 data, point estimates of a selection of statistics measuring the concentration of net worth among U.S. households increased since 1983, and a range of tests provide some evidence that the change may be statistically significant. The point estimate of the Gini coefficient for net worth, a common measure of concentration, rose from 0.777 in 1983 to 0.793 in 1989 according to the model-based weight, or to 0.800 according to the design-based weight. Similarly, the point estimate of the proportion of net worth held by the top one-half percent of the net worth distribution was estimated to have grown from 24.1 percent in

1983 to 28.8 percent in 1989 under the model-based weight, or to 29.1 percent under the design-based weight. Most of the apparent shift to the top one-half percent of the net worth distribution arose because the share of the rest of the top ten percent of the distribution declined.

Although these changes may appear large, formal significance tests cannot be conducted at this time. While we are able to compute estimates of sampling error for the 1989 model-based weights, we are not yet able to make such calculations for the 1983 figures. However, we considered some alternatives to span a range of assumptions about the sampling variance of the 1983 estimate: (1) the variances of the 1983 estimates are equal to 1.41 (the square root of two) times the 1989 variances -- under the assumption that the chief difference between the two samples is the fact that the 1989 list sample is about twice the size of that in 1983 -- and (2) the 1983 variances are equal to twice the 1989 variances. Tests developed under both of these assumptions implied that the change in the Gini coefficient is statistically significant at the five percent level. Similarly, for the proportion of net worth held by the top one-half percent of the net worth distribution, the tests developed implied that the increase in concentration of net worth was statistically significant at the five percent level. However, all of these tests are still limited.

First, we are comparing results at very different stages of the business cycle and it is not known how much of the apparent change might be due to that difference. Second, at this point we are only making allowances in our tests for sampling error, not error attributable to imputation or to other data problems. This omission will tend to overstate the precision of our measurement of change. Third, the 1983 and 1989 sample designs and the weights developed are quite different. The effect of this difference on our estimates is unknown. Fourth, the variance estimates we have used depend on the reliability of the bootstrap method we have employed.

The performance of the bootstrap in this setting is not yet fully understood. Finally, alternative tests may be statistically preferable.

As we have noted, this paper represents a progress report of on-going research on the creation and use of design-based and model-based weights in estimation. In future papers, we plan to refine our procedures. First, we expect that we can construct an alternative design-based weight and corresponding bootstrap sample weights that will incorporate additional information, yet still be closer to classical design-based weights than our model-based weights. Second, we plan to revisit the 1983 SCF design in order to recompute weights based on comparable methodologies and to create corresponding replicate weights for variance estimation. Third, we hope to take explicit account of variances of our estimates that are due to imputation by using the multiply-imputed replicates available in the 1989 SCF. Fourth, we would also like to refine the significance tests we apply to account formally both for the small variation in the earlier estimates of concentration statistics as well as for the differences in the estimates obtained using the two types of weights for 1989. Finally, we would like to compare the survey estimates with aggregate data (particularly the Federal Reserve Flow-of-Funds Accounts, and estate and income data maintained by SOI) as a check on the measurement process.

Appendix: Computation of Model-based Weights

As noted in the text, the first step in the calculation of either the design-based weights or the model-based weights is the determination of the probabilities given in equation one on page eight -- the probabilities of observation of each area-probability case under both the area-probability and list designs, and the probabilities of observation of each list case under the area-probability and list designs. The principal difference in the two types of weights is that the model-based weights incorporate more auxiliary information to shape these first-stage probabilities than do the design-based weights.[33]

The calculation of the probability of observation of an area-probability case under the area-probability design, for which we have no additional systematic frame information, is based on the original design and the response rate attained by geographic region. The original probability of selection of each case is adjusted uniformly within geographic area to compensate for nonresponse. This is the same procedure used in the case of the design-based weight.

However, in sharp contrast to the design-based weight, the calculation of the probability of observation in the list frame of the list sample cases ignores the original SCF sample design almost entirely. The list sample cases were drawn from the SOI 1987 individual tax file, which is itself a sample -- albeit a very large sample -- from the universe of U.S. individual tax-filers.[34] A great deal of information is known about both survey respondents and nonrespondents in the

[33] See Kennickell and Woodburn [1992] for a more complete description of the construction of these weights.

[34] The sample design for the 1987 SOI Individual study can be found in Internal Revenue Service [1990].

list frame.[35] Our strategy is to use this auxiliary information to estimate the representation in the population of each list observation using the entire SOI file to estimate a response propensity model.[36] Formally, we compute an unweighted probit model using the entire SOI file (not just the sample selected for the 1989 SCF) to estimate the conditional probability of observation within the SOI file, given characteristics present in that file.[37] The probability of observation for a list case under the list frame is given by the product of the conditional probability predicted by this model and the probability of selection into the SOI sample. A special adjustment was made in the calculation for observations with negative net worth. The probit model tends to imply unreasonably high weights for these cases. To mitigate this problem, the list cases with negative net worth were assigned the overall median predicted propensity score.

These predicted list weights were post-stratified to ensure that the inverse of the probabilities adds up to the underlying population distribution of the list frame. List cases with negative net worth were excluded from these further adjustments. The weights for list cases drawn from wealth index strata corresponding to a wealth index value greater than $100,000 were controlled to totals estimated from the entire 1987 SOI sample adjusted to represent the

---

[35] Among the very few pieces of survey information that have been connected to the list frame is response status of each sample case. It should be emphasized that while we do know the set of cases that responded, we have no knowledge of which interviews correspond to which administrative records.

[36] Rubin [1985] proposes the use of propensity scores in applied bayesian inferences and discusses the application to adjusting for unit nonresponse. Little and David [1983] and Lepkowski, Kalton, and Kaspryzk [1989] use the propensity score method to adjust for nonresponse in panel survey settings. More recently, Ekholm and Laaksonen [1991] discuss the application of propensity scores in the Finnish Household Budget Survey.

[37] The estimated model is given in Kennickell and Woodburn [1992].

1989 population of families.[38] The control for the lowest wealth index stratum -- wealth index less than $100,000 -- was computed as the total number of families in 1989 less the total of the controls for the wealth index strata with wealth index greater than $100,000 as noted above, less the weighted frequency of cases with negative net worth, and less an estimate of the number of nonfilers.[39] Using the resulting weights, weighted distributions of key income variables in the list frame compared favorably to the corresponding distributions estimated using the entire list frame.[40]

We are able to do more limited modeling of the implied list weights for area-probability cases. Area-probability cases that did not file a 1988 tax return as reported in the interview are assumed not to have filed a 1987 return and, thus, to have a zero list weight. For the remaining cases, those having $100,000 and more of net worth are treated separately. For these wealthier cases, we model the propensity-score-adjusted list weights directly in terms of a number of survey variables. The model was estimated in logs using the list sample cases with net worth of $100,000 and more and used to predict weights for area-probability cases.[41] The predicted weights were adjusted to sum to a population total estimated for this group using the list sample.

[38] In order to estimate the 1989 population of families, two adjustments were made to the estimates computed using the 1987 SOI Individual file. First, the SOI sampling weights of returns for married individuals filing separately were halved. These adjusted weights were used to compute the 1987 population of families. The second adjustment was simply to uniformly adjust these totals in wealth index strata to reflect the rate of growth for all families between 1987 and 1989.

[39] The number of nonfilers was estimated using the survey measure for filing status for tax year 1988 weighted by the nonresponse adjusted area-probability weight.

[40] The distributions were compared visually by the means of q-q distributions. More details are available in Kennickell and Woodburn [1992]. The model-based weights performed much better in this sense than the design-based weights.

[41] A copy of the model is given in Kennickell and Woodburn [1992].

Largely because the list sample is very thin at the bottom of the net worth distribution, estimates of the list weight model that included the entire list sample produced very unstable values for the predicted list weights for lower-wealth families. The remaining area sample cases with less than $100,000 of net worth were assigned the median propensity-score-adjusted list weight for list cases with less than $100,000 of net worth. The weights for these observations were adjusted to sum to the 1989 population total less the estimated number of nonfilers and the number of cases estimated to have $100,000 or more of net worth.

The probability of observation of list sample cases under the area design is assumed to be zero for list cases with $2.5 million or more of net worth and for those with negative net worth.[42] Other list cases are assigned the median area-probability weight of area-probability cases within the same region and MSA type.[43] The probabilities were rescaled so that the implied weights summed to the 1989 population total minus the estimated number of nonfilers and minus the estimated number of observations with $2.5 million or more of net worth estimated using the propensity-score-adjusted list weights.

The component probabilities were merged using the formula given by (1). This formula is the appropriate one to use in the case where the probabilities are known without error; however, in this case, the probabilities are estimates with some bias and variance properties. We have chosen not to implement any explicit adjustments for such uncertainty at this point. Rather, we capture the variance in the computation of the bootstrap weights.

---

[42] This assumption likely understates the probability of observation for higher-wealth observations since there are a small number of relatively wealthy area sample cases. While one might argue for a higher cut-of, such a change has little effect.

[43] Several possible models for these probabilities were investigated. Unfortunately, the models are too sensitive to be useful for weight adjustment.

The combined weights were subjected to post-stratification and raking. First, the weights were adjusted to reproduce exactly the number of families estimated by the list sample to have net worth in the categories $1 million to $2.5 million, $2.5 million to $10 million, and $10 million to $250 million. List and area-probability observations having net worth of $1 million or more and list cases with negative net worth less than - $100,000 were excluded from further adjustments. The remaining observations were raked for three iterations to distributions of age by homeownership and region by MSA status.

Bibliography

AVERY, R.B., ELLIEHAUSEN, G. E., and KENNICKELL, A. B. (1988). "Measuring Wealth with Survey data: An Evaluation of the 1983 Survey of Consumer Finances," *Review of Income and Wealth*, Series 34, No. 4, pp. 339-369.

AVERY, R. B., and KENNICKELL, A. B. (1990). "Measurement of Household Saving Obtained from First-Differencing Wealth Estimates," Paper presented at the Twenty-First General Conference of the International Association for Research in Income and Wealth, Lahnstein, Germany, August, 1989.

AVERY, R. B., and KENNICKELL, A. B. (1992). "Household Saving in the U.S.," *Review of Income and Wealth*, Series 37, Number , pp. 409-432.

EKHOLM, A. and LAAKSONEN, S. (1991). "Weighting via Response Modeling in the Finnish Household Budget Survey," *Journal of Official Statistics*, 7, 3, 1991, pp. 325-337.

HARTLEY, H. O. (1962). "Multiple Frame Surveys," *Proceedings of the Social Statistics Section*, American Statistical Association, pp. 203-206.

HEERINGA, S., JUSTER, F. T., and WOODBURN, R. L. (1991). "The 1989 Survey of Consumer Finances: A Survey Design for Wealth Estimation," *Statistics of Income and Related Record Research*, 1990-1991, Internal Revenue Service (forthcoming).

HEERINGA, S., and WOODBURN, R. L. (1992). "The 1989 Survey of Consumer Finances, Sample Design Documentation," (forthcoming).

HOLT, D., and SMITH, T. M. F., (1979). "Post-Stratification," *Journal of the Royal Statistical Society*, A, 142, Part 1, pp. 33-46.

HUGGINS, V. J., and FAY, R. E., (1988). "Use of Administrative Data in SIPP Longitudinal Estimation," *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

INTERNAL REVENUE SERVICE, (1990). *Individual Income Tax Returns 1987*, Department of the Treasury, pp. 13-17.

JOHNSON, B. W., SCHWARTZ, M. ?., (1990). "Estate Tax Returns," *Statistics of Income Bulletin*, Spring.

JOHNSON, B. W., and WOODBURN, R. L. (1992). "Evaluation of the Estate Multiplier Methodology and Comparison with Survey Wealth Estimates: Some Results for 1982 and 1989," (forthcoming).

KENNICKELL, A. B., and SHACK-MARQUEZ, J. (1992). "Changes in Family Finances from 1983 to 1989: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, Board of Governors of the Federal Reserve System, 78, No. 1, 1992, pp. 1-18.

KENNICKELL, A. B. (1991a). "Imputation of the 1989 Survey of Consumer Finances: Stochastic Relaxation and Multiple Imputation," *Proceedings of the Section on Survey Research Methods*, American Statistical Association (forthcoming).

KENNICKELL, A. B. (1991b). "Technical Codebook for the 1989 Survey of Consumer Finances," Board of Governors of the Federal Reserve System.

KENNICKELL, A. B., and WOODBURN, R. L. (1992). "Development of Model-Based Weights for the 1989 Survey of Consumer Finances," Technical Documentation. (forthcoming).

LEPKOWSKI, J., KALTON, G. AND KASPRYZK, D. (1989). "Weighting Adjustments for Partial Nonresponse in the 1984 SIPP Panel," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 296-301.

LITTLE, R. J. A. (1991). "Post-stratification," *Journal of the American Statistical Association* 1992 or Proceedings 1991.

LITTLE, R. J. A. and DAVID, M. H. (1983). "Weighting Adjustments for Nonresponse in Panel Surveys," Working Paper. Washington D.C., U.S. Bureau of the Census.

OH, H. L., and SCHEUREN, F. J. (1978). "Multivariate Raking Ratio Estimation in the 1973 Exact Match Study," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 107-111.

PROJECTOR, D. S. AND WEISS, G. S. (1966). *Survey of Financial Characteristics of Consumers*, Board of Governors of the Federal Reserve System, Washington, DC.

RAO, J. N. K., and WU, C. F. J. (1988). "Resampling Inference With Complex Survey Data," *Journal of the American Statistical Association*, 83, 401, pp. 231-241.

RUBIN, D. B. (1985). "The Use of Propensity Scores in Applied Bayesian Inference," *Bayesian Statistics*, 2, pp. 463-472.

SKINNER, C. J., HOLT, D., and SMITH, T. M. F. (editors) (1989). *Analysis of Complex Surveys*, John Wiley and Sons.

WILSON, O., and SMITH, W. J. Jr. (1983). "Access to Tax Records for Statistical Purposes," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 595-601.

WOODBURN, R. L. (1991). "Using Auxiliary Information to Investigate Nonresponse Bias," *Proceedings of the Section on Survey Research Methods,* American Statistical Association.

Table 1: Holdings of net worth and its components (billions of current dollars) and shares of total, by various percentile groups of the distribution of net worth; using model-based weights,1989.

	Percentile of the net worth distribution									
	All households		*0 to 89.9*		*90 to 99*		*99 to 99.5*		*99.5 to 100*	
Item	Holdings	% of total	Holdings	% of total	Holdings	% of total	Holdings	% of total	Holdings	% of total
Assets	17,974.5	100.0	6,611.7	36.8	5,346.3	29.8	1,334.0	7.4	4,682.5	26.0
	704.7		*161.7*	*1.7*	*295.80*	*1.3*	*180.7*	*0.6*	*493.0*	*2.0*
Princ. residence	5,803.7	100.0	3,775.5	65.1	1530.8	26.4	190.1	3.3	307.4	5.3
	117.3		*99.8*	*1.2*	*99.10*	*1.5*	*27.3*	*0.5*	*23.5*	*0.4*
Other real estate	2,491.3	100.0	493.3	19.9	882.4	35.5	281.6	11.4	834.1	33.4
	189.2		*35.3*	*1.3*	*111.10*	*3.7*	*62.6*	*2.5*	*150.2*	*4.5*
Stocks	865.4	100.0	126.5	14.6	311.4	35.9	100.2	11.5	327.3	37.9
	50.9		*17.5*	*1.9*	*34.70*	*2.9*	*28.2*	*2.9*	*27.4*	*3.3*
Bonds	1,299.5	100.0	72.0	5.9	203.0	16.1	146.2	11.8	678.2	66.1
	321.8		*15.1*	*2.0*	*30.30*	*2.8*	*58.6*	*5.1*	*305.7*	*8.1*
Trusts	495.2	100.0	35.0	7.2	203.3	41.2	56.4	11.3	200.6	40.2
	62.9		*5.7*	*2.0*	*54.90*	*10.1*	*24.0*	*4.6*	*68.1*	*11.3*
Life Insurance	256.9	100.0	155.3	60.5	69.3	26.9	13.3	5.1	19.0	7.4
	19.3		*10.9*	*2.9*	*10.60*	*3.1*	*4.5*	*1.6*	*1.5*	*0.5*
Checking accts	202.0	100.0	95.8	47.7	57.1	28.5	12.5	6.1	36.6	17.7
	22.0		*7.5*	*3.4*	*7.60*	*4.2*	*9.0*	*4.2*	*17.5*	*7.5*
Thrift accounts	383.6	100.0	138.8	36.2	188.6	49.1	34.6	9.0	21.5	5.6
	24.0		*14.3*	*2.9*	*15.20*	*1.8*	*7.7*	*1.9*	*6.4*	*1.7*
Other accounts	1,680.4	100.0	747.5	44.5	627.9	37.4	112.0	6.7	193.0	11.5
	47.6		*40.2*	*2.5*	*40.90*	*2.0*	*25.1*	*1.5*	*33.5*	*1.9*
Businesses	2,969.6	100.0	267.2	9.1	864.1	29.1	310.7	10.2	1,527.6	51.6
	334.9		*24.5*	*1.3*	*134.80*	*3.2*	*162.6*	*4.1*	*16.6*	*4.8*
Automobiles	704.5	100.0	532.0	75.5	134.9	19.1	14.3	2.0	23.4	3.3
	14.5		*8.5*	*1.4*	*10.10*	*1.2*	*1.6*	*0.2*	*4.6*	*0.7*
Other assets	822.3	100.0	172.7	21.6	273.4	34.0	62.2	7.7	313.9	36.7
	137.0		*7.5*	*4.1*	*63.60*	*8.7*	*15.2*	*2.1*	*144.2*	*12.9*
Liabilities	2,825.2	100.0	1,805.5	63.9	619.1	21.9	75.1	2.7	325.4	11.5
	78.0		*85.6*	*2.0*	*59.10*	*2.1*	*18.9*	*0.7*	*56.7*	*1.8*
Princ. res. debt	1,623.4	100.0	1,236.6	76.3	322.6	19.9	25.7	1.6	35.5	2.2
	55.8		*54.3*	*2.2*	*34.40*	*1.9*	*5.5*	*0.4*	*5.4*	*0.3*
Other r/e debt	605.6	100.0	129.3	21.4	198.9	33.1	42.0	6.9	235.4	38.7
	54.8		*18.6*	*2.4*	*22.50*	*4.7*	*17.7*	*2.6*	*41.1*	*3.7*
Other debt	596.2	100.0	437.7	73.5	96.6	16.2	7.4	1.3	54.5	9.1
	24.6		*17.5*	*2.7*	*13.40*	*2.4*	*2.0*	*0.4*	*19.6*	*2.8*
Net worth	15,149.3	100.0	4,806.1	31.8	4,727.2	31.2	1,259.0	8.3	4,357.1	28.8
	649.5		*163.6*	*1.8*	*262.30*	*1.3*	*182.9*	*0.9*	*441.1*	*2.0*
Total Family Income	3,144.,1	100.0	2,187.9	69.6	614.8	19.5	110.1	3.5	231.3	7.4
	67.9		*49.0*	*1.2*	*33.40*	*0.9*	*17.1*	*0.5*	*49.6*	*1.5*
Memo Items:										
Number of obs.	3,143.0	100.0	2,056.0	65.4	639.0	20.3	121.0	3.8	327.0	10.4
Num. families (mil.)	93.1	100.0	83.8	90.0	8.4	9.0	0.5	0.5	0.5	0.5

Standard errors due to imputation and sampling are given in italics.

Table 2: Holdings of net worth and its components (billions of current dollars) and shares of total, by various percentile groups of the distribution of net worth; using design-based weights; 1989.

	Percentile of the net worth distribution									
	All households		*0 to 89.9*		*90 to 99*		*99 to 99.5*		*99.5 to 100*	
Item	Holdings	% of total	Holdings	% of total	Holdings	% of total	Holdings	% of total	Holdings	% of total
Assets	20,092.7	100.0	7,135.2	35.5	6,355.3	31.6	1,287.8	6.4	5,314.4	26.4
Princ. residence	6,415.5	100.0	3,993.5	62.2	1,882.2	29.3	217.3	3.4	322.5	5.0
Other real estate	2,907.2	100.0	519.8	17.9	1,136.1	39.1	259.6	8.9	991.7	34.1
Stocks	1,062.9	100.0	184.7	17.4	404.5	38.1	79.7	7.5	394.1	37.1
Bonds	994.6	100.0	65.4	6.6	281.1	28.3	57.7	5.8	590.3	59.4
Trusts	413.9	100.0	54.8	13.2	137.3	33.2	119.3	28.8	102.6	24.8
Life Insurance	359.8	100.0	206.7	57.4	103.2	28.7	21.2	5.9	28.7	8.0
Checking accts	210.0	100.0	111.7	53.2	58.2	27.7	8.7	4.1	31.3	14.9
Thrift accounts	452.5	100.0	165.8	36.6	208.5	46.1	31.8	7.0	46.4	10.3
Other accounts	1,968.8	100.0	807.4	41.0	755.1	38.4	134.9	6.9	271.5	13.8
Businesses	3,718.9	100.0	264.4	7.1	939.2	25.3	285.3	7.7	2,230.0	60.0
Automobiles	773.5	100.0	557.7	72.1	157.1	20.3	13.9	1.8	44.7	5.8
Other assets	815.0	100.0	203.2	24.9	293.0	36.0	58.3	7.2	260.5	32.0
Liabilities	3,129.2	100.0	1,923.4	61.5	744.8	23.8	83.0	2.7	378.0	12.1
Princ. res. debt	1,749.1	100.0	1,331.8	76.1	350.3	20.0	28.6	1.6	38.4	2.2
Other r/e debt	764.4	100.0	140.0	18.3	302.2	39.5	44.7	5.8	277.5	36.3
Other debt	615.7	100.0	451.6	73.3	92.3	15.0	9.7	1.6	62.1	10.1
Net Worth	16,963.5	100.0	5,211.7	30.7	5,610.6	33.1	1,204.8	7.1	4,936.4	29.1
Total Family Income	3,623.2	100.0	2,389.2	65.9	730.9	20.2	130.7	3.6	372.4	10.3
Memo Items:										
Number of obs.	3,143.0	100.0	2113.0	67.2	584.0	18.6	118.0	3.8	328.0	10.4
Num. families (mil.)	93.1	100.0	83.8	90.0	8.4	9.0	0.5	0.5	0.5	0.5

Table 3: Holdings of net worth and its components (billions of current dollars) and shares of total, by various percentile groups of the distribution of net worth; 1983.

	Percentile of the net worth distribution									
	All households		*0 to 89.9*		*90 to 99*		*99 to 99.5*		*99.5 to 100*	
Item	Holdings	% of total	Holdings	% of total	Holdings	% of total	Holdings	% of total	Holdings	% of total
Assets	11,708.9	100.0	4,432.4	37.9	3,895.7	33.3	795.0	6.8	2,585.8	22.1
Princ. residence	3,738.5	100.0	2,448.0	65.5	977.7	26.2	140.2	3.8	172.6	4.6
Other real estate	1,687.5	100.0	405.2	24.0	668.1	39.6	116.6	6.9	497.6	29.5
Stocks	1,041.1	100.0	105.9	10.2	335.4	32.2	83.4	8.0	516.4	49.6
Bonds	367.0	100.0	22.6	6.2	157.1	42.8	29.8	8.1	157.6	42.9
Trusts	309.4	100.0	25.0	8.1	78.9	25.5	17.5	5.7	187.9	60.7
Life Insurance	284.8	100.0	175.8	61.7	62.5	21.9	12.3	4.3	34.2	12.0
Checking accts	119.4	100.0	66.0	55.3	36.9	30.9	6.5	5.4	10.1	8.5
Thrift accounts	154.3	100.0	47.4	30.7	74.2	48.1	17.2	11.1	15.4	10.0
Other accounts	1,049.1	100.0	517.5	49.3	384.0	36.6	46.6	4.4	101.0	9.6
Businesses	2,284.3	100.0	215.5	9.4	925.3	40.5	294.9	12.9	848.7	37.2
Automobiles	373.3	100.0	295.8	79.2	65.5	17.5	5.4	1.4	6.7	1.8
Other assets	300.1	100.0	107.5	35.8	130.2	43.4	24.7	8.2	37.6	12.5
Liabilities	1,507.6	100.0	1,030.3	68.3	294.9	19.6	58.8	3.9	123.6	8.2
Princ. res. debt	864.6	100.0	691.8	80.0	137.7	15.9	21.0	2.4	14.1	1.6
Other r/e debt	324.1	100.0	117.5	36.3	107.7	33.2	15.8	4.9	83.0	25.6
Other debt	318.9	100.0	221.0	69.3	49.5	15.5	22.0	6.9	26.4	8.3
Net worth	10,201.3	100.0	3,402.1	33.3	3,600.8	35.3	736.2	7.2	2,462.2	24.1
Total Family Income	2,254.1	100.0	1,630.4	72.3	431.1	19.1	61.4	2.7	131.3	5.8
Memo Items:										
Number of obs.	4,103.0	100	3,340.0	81.4	477.0	11.6	8.2	2.0	204.0	5.0
Num. families (mil.)	83.9	100	75.5	90.0	7.6	9.0	0.4	0.5	0.4	0.5