Modeling Wealth with Multiple Observations of Income: Redesign of the Sample for the 2001 Survey of Consumer Finances[1]

Arthur B. Kennickell
Federal Reserve Board, Mail Stop 153, Washington, DC 20551

KEY WORDS: Sampling, wealth, income, SCF

This paper reports on research that underlies the redesign of a key part of the sample for the Survey of Consumer Finances (SCF) for the 2001 wave of the survey, and it documents the implementation of the new design. This work builds on the history of sampling research for the survey dating back to the 1983 wave.

The sample for the SCF includes an oversample of relatively wealthy families. The stratified design is based on a mapping from observations of components of income to an estimate of wealth. The mapping is imperfect, and three imperfections seem particularly important: First, at any given time, rates of return that connect assets with capital income may vary widely across individuals depending on luck, information, and local economic conditions. Second, some assets, such as 401(k) accounts, do not generate regularly observable returns. Finally, transitory factors, such as the timing of income receipts or unusually good luck, may cause the income that is observed in a given period to have a noisy relationship to the underlying assets that generate those returns. It is very difficult to address the first two concerns directly. However, the third can be addressed by using multiple observations of income to model wealth; an effort to do so is the principal focus here.

The first section of this paper gives a brief overview of the SCF and the sample design. The next section provides motivation for a reexamination the sample design. The third section describes the implementation of the redesign of the sample for the 2001 SCF. A final section discusses areas for future research.

I. Overview of the SCF and Sample design

The SCF is undertaken every three years by the Federal Reserve Board (FRB) primarily as a study of household wealth and use of financial services.[2] To this end, the subject matter of the questionnaire focuses on detailed components of assets and liabilities and supporting information.

Because household wealth in the U.S. is highly concentrated (Kennickell, 2000b), a very large sample would be required to make reliable wealth estimates without some type of oversampling by wealth. At the same time, the problem of adequate representation of wealth is amplified by the nonrandom nature of nonresponse to the survey. Although many factors enter into nonresponse in the SCF (Kennickell, 1999b), it is clear that wealthier respondents are less likely to participate. Lower participation may be driven by a greater sensitivity about privacy issues, a greater perceived value of the time required for the interview, the greater difficulty interviewers face in getting beyond gatekeepers in order to request participation, or other factors. The net result is that without some means of addressing the nonrandom nonresponse, many of the survey estimates would be severely biased.

The SCF sample is a dual-frame design. One part is a multi-stage national area-probability design (Tourangeau et al., 1993), which provides good coverage of the general population. This sample is selected by the National Opinion Research Center at the University of Chicago, the contractor for data collection since 1992.

The second part is a list sample, which is selected by FRB staff from statistical records derived from tax returns. These records are provided by the Statistics of Income Division (SOI) of the Internal Revenue Service to the FRB under strict controls on the use of the data. This file is generated from a sample from the full set of tax returns filed, and it is specially edited using procedures designed to yield data to support tax research at the Office of Tax Analysis and the Joint Economic Committee (see IRS, 1998). Since the 1989 survey, the SOI data used for the SCF sample have come from the tax year (generally, the year in which the reported income was earned) two years before the date of the survey. The list sample is stratified using a "wealth index" computed using income data in order to predict a rank ordering of people by wealth. This stratification allows both oversampling of wealthy families and targeted nonresponse adjustment. The area-probability and list samples are pooled using weights that are designed to respect the relative strengths of each part of the sample (Kennickell and Woodburn, 1999).

Because the wealth index used in the list sample stratification has such a powerful effect on the estimation efficiency of the SCF, refinement of the index has always been an important part of the methodological research supporting the survey (see Kennickell 1999a for a history). The list sample for the 1983 survey, the first of the series, was stratified by income categories created using a set of rules that have not been cleared for public release. The 1989 SCF was the first to use a wealth index approach, and the construction of that index is discussed in Heeringa *et al.* (1994). In its pure form, this general type of index for a given time might be expressed as

$$WINDEX0 = \sum_j \frac{1}{r_{ij}} Y_{ij} ,$$

where "i" indexes individuals, "j" indexes components of

capital income, and r_{ij} is the rate of return associated with capital income component Y_{ij}. Thus, if a person had interest income of $100 and the associated rate of return were 5 percent, then the contribution of this asset to the wealth index would be $2,000.

One problem in implementing such an index is that no direct information is available *a priori* that would allow the use of person-specific rates of return. Thus, as WINDEX0 has been implemented, a common rate of return is assumed for all individuals. Another problems is that not all assets generate regularly observable returns (Kennickell and McManus, 1993). For example, the only income associated with owned personal residences appears when such properties are sold; the returns on 401(k) accounts, IRAs, and Keogh accounts only appear as income when funds are withdrawn. Some items may only appear as entries on estate tax returns. An index of this broad form was used in selecting the list samples for the 1989 and 1992 surveys.

Beginning with the 1995 SCF, the wealth index used for stratification was altered to include a component calculated from a model estimated in a regression of actual wealth observed in the previous survey on its corresponding original frame data. This model, referred to as "WINDEX1", offers a flexible way of accounting for systematic patterns in the structure of various types of income and other observed factors on wealth. However, the underlying structure may change as rates of return change over time.

To hedge against misclassification from the WINDEX1 model, the 1995 and later surveys have pooled estimates of WINDEX0 and WINDEX1; that pooled estimate is referred to as "WINDEXM". For the pooling, the two indices are adjusted to have the same median and inter-quartile range.

Figure 1 shows the distribution of the pooled wealth index by the unweighted deciles of the net worth of the portion of the SCF list sample that was interviewed in 1998.[3] Ideally, the figure would show strong clustering of the mass of each distribution along the diagonal from the lower left-hand corner to the upper right-hand corner. Although it is clear that WINDEXM does do a reasonably good job of discriminating between very high wealth and very low wealth, across the middle of the distribution the power of the model is lower. It is clear that there is room for improvement.

II. Motivation for redesign of the list sample

Despite the efforts described above, the WINDEXM model is only a sketch of the deeper structural model of the connections between wealth and income, which should include modeling of portfolio decisions, employment choices, other planning decisions, individual effects of the larger economy, etc. Idiosyncratic factors are also very likely to be important. Although the SOI data offer only limited scope for identifying additional structural and idiosyncratic factors directly, there is still a prospect that better proxies might yield a more reliable model for indexing wealth.

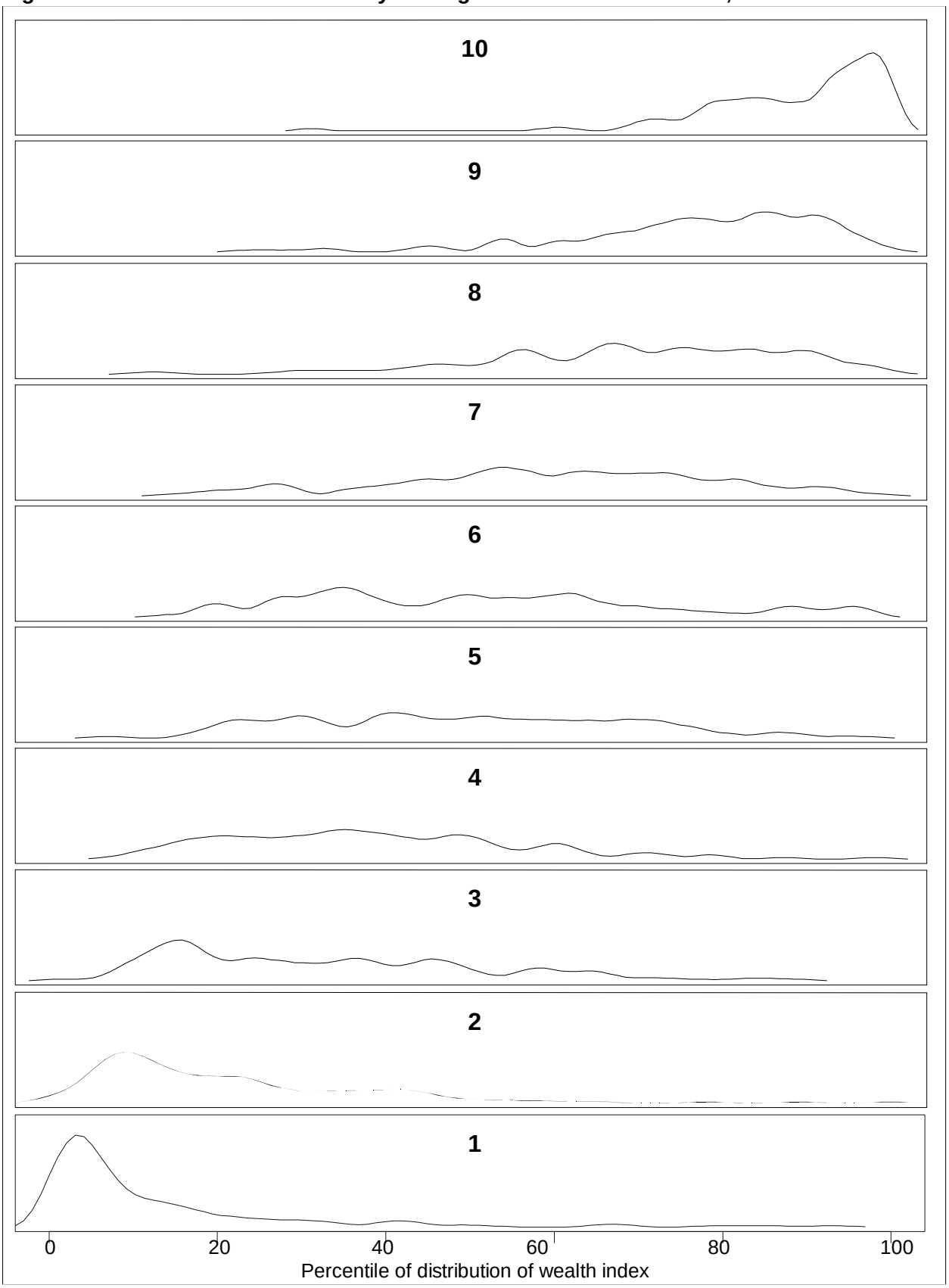
One dimension of the data that has previously not been exploited for the SCF sample is time. Income may vary over time for a number of reasons. Some income variations may parallel changes in the valuation of the underlying assets that generate the incomes. Using cross-sectional data for the sample from two tax years before the survey poses a risk of over- or under- stating wealth at the time of the survey. But if positive and negative shifts are equally likely, then pooling multiple years of income would damp such shifts without introducing net bias in estimated wealth rankings.

Other changes in income may be less strongly related to wealth. Some people may be able to time the receipt of their incomes for particular needs or for an advantage in their taxes. Others may have assets that have a relatively spikey payout profile. Many people are affected by changes in the larger economy, through their employment and wages, their interest earnings, and other such factors, and these events may affect both asset holdings and returns of households differently over time. Averaging over observations of the same unit over time would have the effect of smoothing out such transitory fluctuations in income.

Although extensive time series of data are available in principle, practical obstacles to obtaining the necessary data for sampling are large. First, access is very limited. Although the FRB has a contract with SOI that allows the use of IRS data, every instance of use must be separately justified. Given the abstract potential for abuse, it would be difficult to argue for substantially easier access. Arranging for access to even the types of records used previously for the SCF requires substantial time from many people. Second, it also requires time to process the data needed for the sample. Individuals mostly file tax returns by April following the tax year, but some people, most often those with particularly complex tax returns, may request an extension and file later. The SOI file that serves as the basis of the SCF list sample requires substantial editing. Some of this editing corrects errors and reclassifies the reported data in a way that is more appropriate for tax research. Although "advance" data for a given tax year are typically available by around December of the year following the tax year, the final data have never been available for the SCF sooner than April two years after the tax year. Thus, the SCF sample can be based on no more recent data than such a file. Third, the SOI file is itself a sample, and observations are not necessarily retained from year to year. Thus, linking individual records would require access to IRS master file data, a more restricted source of information.

For the design of the 2001 SCF list sample, extraordinary efforts were made by SOI staff to make it possible to use both the type of SOI data file used in the past, as well as IRS master file data. Four sets of data

Figure 1: Distribution of WINDEXM by unweighted deciles of net worth, 1998 SCF.



were provided: master file data for tax years 1996 and 1997 for cases corresponding to the full original list sample in the 1998 SCF, the edited SOI file for tax year 1999, and a set of master file data for tax year 1997 for observations in the full 1999 SOI file. The 1999 SOI file and the 1997 master file data are the most recent files of these types that are available. In order to gauge the costs and benefits of using these data sources together, it is important to assess the nature of the SOI editing process as well as the variability of income measures over time. To this end, the master file and SOI data for 1997 can be compared to assess the effects of mixing data from the two files, and the 1996 SOI data can be compared with the 1997 master file data to assess the variability of income over time.

Figures 2a and 2b show various statistics on the variation in measurements induced by the editing of the 1996 SOI data, the basis of the 1998 SCF list sample, conditioned on WINDEX0 computed from the SOI data. In each plot, the solid line indicates the percent of tax filers who had a change in the value of the item as a result of editing, the dashed line indicates the mean percent change relative to the SOI value, and the dotted line shows the corresponding standard deviation of the percentage change. The distribution of adjusted gross income (AGI) is different for only a small fraction of observations away from the very top of the WINDEX0 distribution, and the mean and standard deviation of the change is small across the distribution. At the same time, variations in the percent of observations with changes in the components of AGI (not shown) are substantially larger. With increasing values of WINDEX0, wage and salary income shows an increasingly greater value in the SOI data, largely reflecting what appears to be a systematic reclassification of S-corporation and farm income and Schedule C income to wages and salaries. The net effect of editing changes on WINDEX0 is to induce a large number of relatively small changes across the distribution of WINDEX0 computed with the SOI data. The standard deviation of the percent change is roughly 10 percent across the range of WINDEX0. However, the critical question is how large these changes are relative to the changes in the data over time.

Figures 3a and 3b address the magnitude of the

Figure 2: Differences between 1996 SOI and master file data; percent having any difference, mean percent change, standard deviation of percent change; by WINDEX0; 1998 SCF sample.

Figure 2a: Adjusted gross income.

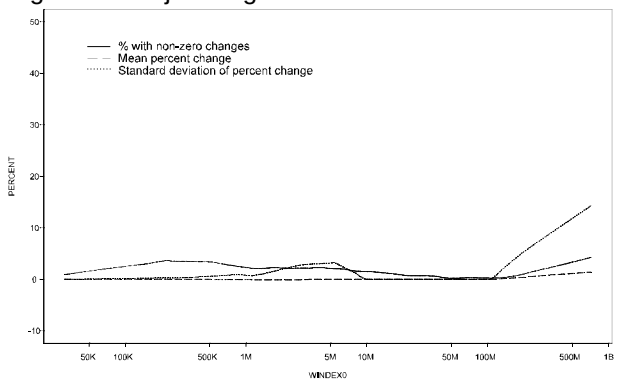


Figure 2b: WINDEX0.

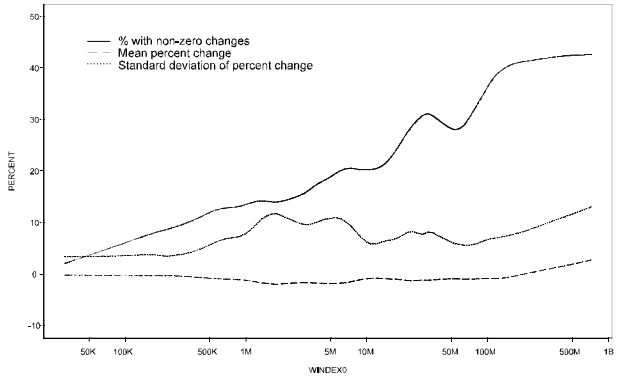


Figure 3: Differences between 1997 master file data and 1996 SOI data; percent having any item in either data, mean percent change (in constant dollars), standard deviation of percent change; by WINDEX0; 1998 SCF sample.

Figure 3a: Adjusted gross income.

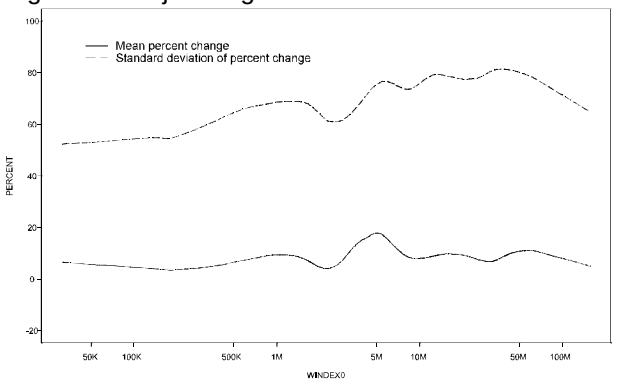
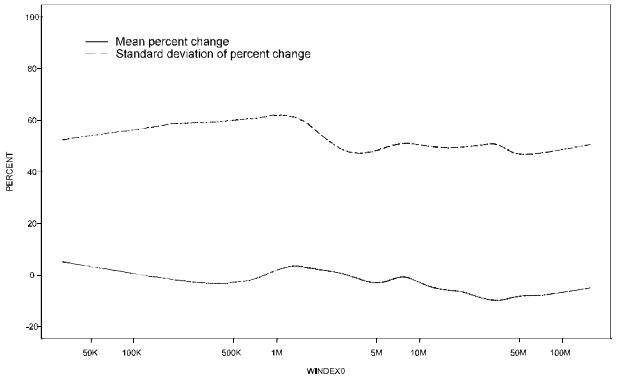


Figure 3b: WINDEX0.



changes over time in the same variables. The solid lines show the mean percent change from 1996 to 1997, and the dashed lines gives the standard deviations of the percentage changes. All values were computed using the same data and assumptions with an additional adjustment to put the changes into constant-dollar terms, and observations that had a change in filing status are excluded. The difference in AGI over the period is substantially greater across the entire range of WINDEX0 than the difference in the corresponding comparison of SOI and master file data. By income components (not shown), the largest mean percent differences overall are in interest and dividends; the mean difference there is roughly 10 percent on average, and the standard deviation is about 70 percent. Mean differences in wage and salaries, S-corporation and farm income, and Schedule C income are fairly small, but the standard deviations are quite large. The net effect on WINDEX0 is relatively small changes on average, but a standard deviation of about 70 percent overall. The variation in WINDEX0 is clearly much larger than what one would expect purely from the reclassification error seen in figure 2b. Thus, the data suggest that there may be more to be gained in sampling efficiency by accounting for intertemporal variability in income than would be lost from the noise introduced by using edited and unedited files together.

III. Implementation of the redesign for the 2001 SCF

Given that the data examined suggest that there is a potential gain in sampling efficiency to be had by using the 1997 master file data together with the 1999 SOI data for the design of the 2001 SCF sample, there are two obvious classes of strategies for using these data.

First, one might use all of the variables in the two observations separately to model wealth. If there were a sufficient basis for modeling the mechanism that underlies the observed income variability, the increased number of conditioning variables might yield an improvement in the ability to classify observations in terms of their wealth. But variability in income reported on tax returns is probably mainly a function of two factors: real events ("good/bad luck," unemployment, etc.) and tax considerations. The available tax-based data offer little systematic help in understanding the real factors, and aggregate data from other sources are not useful in understanding individual variations. There is a rich literature that attempts to understand how people manage their finances in response to tax laws, but application of the insights of that literature to the current problem would involve a high degree of speculation. Still, even a non-structural model, if sufficiently agnostic in its construction, might yield improvement from the increased degrees of freedom. However, the practical constraint imposed by the very limited time available for the actual selection of the 2001 sample after the data became available made it impossible to consider this option seriously. Moreover, there is a good argument to be made for altering so critical an element of the SCF in well-defined steps.

The second type of strategy is to use the information from the two periods solely for the purpose of smoothing estimates. While this is a more conservative approach in some ways, it is also probably more immune to modeling errors. Motivated both by caution and the force of time pressures, a version of this strategy was applied to create a smoothed estimate of WINDEXM for the 2001 list sample, following as closely as possible the construction of the 1998 sample (Kennickell, 1998).

The population structure implied by the 1999 SOI file was taken as the point of reference for the sample design. Thus, it was necessary to create a file of matched data from the 1997 master file and for each observation in the 1999 SOI file. However, some people did not file a tax return in both years. In addition, to avoid spurious variability in income measures, it was necessary to exclude cases from the match where there was a difference in tax filing status in the two years. Of the 172,852 observations in the 1999 SOI file, 149,148 could be matched with 1997 master file data.

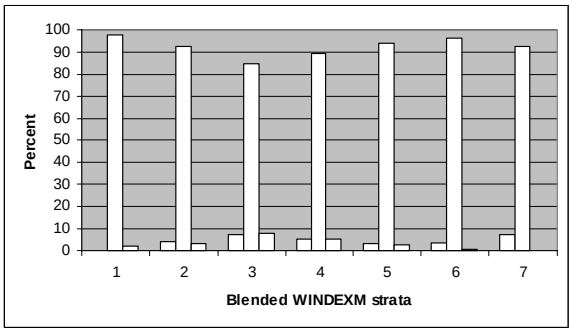
It was necessary to compute an estimate of WINDEXM for every observation in the final data file. Using appropriately updated coefficients, estimates of WINDEX0 were made for every observation for each available year. For the set of observations where both years of data were available, the 1997 estimate of WINDEX0 was adjusted to have the same mean as the 1999 estimate of WINDEX0 for the same set of observations. The smoothed estimate of WINDEX0 was computed as the simple average of the two estimates where both were available, and the 1999 estimate for the other cases. For WINDEX1, the procedure was similar. The blended estimate of WINDEXM is the average of estimate of WINDEX0 and WINDEX1 adjusted to have the same median and inter-quartile range.

The resulting index was used stratify the sample in terms of weighted percentiles of the distribution of WINDEXM. Although exact breaks in this distribution cannot be revealed publicly, the lowest stratum (1) accounts for about 75 percent of the population of tax filers, and the five highest strata together account for about two percent of the population considered.

The results of this stage of the sample construction were compared under the new and old designs. This comparison is motivated principally by two concerns. First, radical differences would be likely to be a result of an error. Second, a central motivation for the redesign is to improve the classification of units across the middle of the distribution into wealth groups, so there should be notable differences in that construction were compared extensively under the new regieme to justify the difficulty and risk of implementing a new procedure.

When the strata were computed using both methods, fewer than 0.15 percent of all observations were

Figure 4: Distribution of unblended stratum assignment around blended assignment.



assigned to a stratum computed under the old method that differed by more than one from that implied by the new method. Figure 4 shows the proportion of observations assigned values under the new method above and below the value under the old method. The tall central bars represent the percent assigned to the same stratum under both methods; the bars to the left (right) represent cases assigned to a lower (higher) stratum under the old method. The differences are generally small. The largest difference occurs in stratum 3, where about 15 percent of cases are assigned a different stratum under the old method. For stratum 1, only 2 percent of cases have a different classification under the old method, while for the top stratum the figure is about 5 percent. The classification based on the blended estimate uses more information than that based on the unblended estimate, and the underlying model framework is the same. Thus, there is no reason to expect the blended estimate to be worse, and there is some *a priori* reason to think that it might be better.

IV. Future research

Upon the completion of data processing, a high priority will be an evaluation of the new sample design. If a significant improvement can be detected, it may justify the effort to gain permission to incorporate a longer series of income data or more complex modeling into the estimation of the principal stratifier for future surveys. At the same time, work must continue to deal with misclassifications resulting from other sources of error, principally interviews with the wrong person (either because of problems in identifying the correct respondent or because of incorrect use of proxies), incorrect reporting or recording of information, and insincere interviews.

Bibliography

Heeringa, S.G., J.H. Conner and R.L. Woodburn [1994] "The 1989 SCF Sample Design and Weighting Documentation," working paper, Survey Research Center, U. of Michigan, Ann Arbor, MI.

IRS [1998] "Individual Income Tax Returns."

Kennickell, Arthur B. [2000a] "Wealth Measurement in the SCF: Methodology and Directions for Future Research," paper presented at the 2000 annual meetings of AAPOR, Portland, May.

_____ [2000b] "An Examination of Changes in the Distribution of Wealth from 1989 to 1998: Evidence from the SCF," Jerome Levy Economics Institute working paper 307.

_____ [2000c] "Revisions to the Variance Estimation Procedure for the SCF" working paper, Board of Governors of the Federal Reserve System.

_____ [1999a] "Using Income Data to Predict Wealth," paper presented at the Annual Meetings of ASSA, New York, 1999b.

_____ [1999b] "Analysis of Nonresponse Effects in the 1995 SCF,"*Journal of Official Statistics*, v. 15 no. 2, 1999a, pp. 283-304.

_____ [1998] "List Sample Design for the 1998 SCF," working paper, Board of Governors of the Federal Reserve System.

_____, M. Starr-McCluer and B.J. Surette [2000] "Recent Changes in Family Finances: Results from the 1998 Survey of Consumer Finances," *Federal Reserve Bulletin*, (January) pp. 1-29.

_____ and R.L. Woodburn [1999] "Consistent Weight Design for the 1989, 1992, and 1995 SCFs, and the Distribution of Wealth," Review of Income and Wealth, pp. 193-215.

_____ and D. McManus [1993] "Sampling for Household Financial Characteristics Using Frame Information on Past Income," *Proceedings of Section on Survey Research Methods*, 1993 Annual Meetings of the ASA, San Francisco, CA.

Tourangeau, R., R.A. Johnson, Jiahe Qian, Hee-Choon Shin, and M.R. Frankel [1993] "Selection of NORC's 1990 National Sample," working paper, NORC at the University of Chicago, Chicago, IL.

Endnotes

1. An unabridged version of the paper is available at http://www.federalreserve.gov/pubs/oss/oss2/method.html. The author is grateful to Ryan Bledsoe for research assistance. The work reported in this paper would not have been possible with the cooperation and efforts of a number of people at the Statistics of Income Division of the IRS, including Nick Greenia, Tom Petska, Michele Rhone, Mike Strudler, and particularly Barry Johnson. James Nunns at the Office of Tax Analysis at the Dept. of the Treasury provided essential support. The opinions expressed in this paper are those of the author alone and they do not necessarily reflect the views of the Federal Reserve Board, the Internal Revenue Service, or the Department of the Treasury.

2. See Kennickell *et al.* (2000) and Kennickell (2000a) for an overview of the survey and it methodology.

3. The list sample is strongly weighted toward the upper end of the wealth distribution, but the exact proportions are withheld for reasons related to disclosure limitation.