The Underlying Methodology of the Estate Multiplier Technique: Recent Improvements for 1989

Barry W. Johnson and R. Louise Woodburn
Internal Revenue Service

The distribution and composition of personal wealth in the United States has attracted considerable attention in recent years. The effects of changes in tax laws, and other public policies, on the economy and on the concentration of wealth, have been widely debated. Accurate and comprehensive estimates of wealth, however, are difficult to obtain because individuals are not generally required to report wealth information on any tax return or other public document. The most common methods of estimating personal wealth are collecting wealth data directly via a survey, such as the Federal Reserve Board's Survey of Consumer Finances [Kennickell and Shack-Marquez, 1992, Avery and Kennickell, 1992], capitalizing income flows reported on individual tax returns [Greenwood, 1981], or using wealth data reported on Federal estate tax returns. This paper focuses on the third alternative, using the 'estate multiplier technique' to estimate the wealth of the living population from estate tax return data collected by the Internal Revenue Service's Statistics of Income Division (SOI).

This paper is divided into four sections. The first will review the estate multiplier technique, the underlying data base, and some of the conceptual challenges associated with the methodology. Next, we will look in detail at the application of the methodology and at the uncertainty associated with the estimation technique by quantifying the effects of our assumptions on the variance of the resulting estimates. We will then evaluate our estimates by comparing them to those developed using the Federal Reserve Board's 1989 Survey of Consumer Finances. Finally, some areas for future research will be discussed.

■ The Technique, Data, and Conceptual Challenges

Researchers have been using the estate multiplier technique since the beginning of the 20th century to draw conclusions about the wealth of the living population by studying the wealth of the deceased. The multiplier technique assumes that estate tax returns, taken as a whole, represent a random sample of the living wealthy population and, thus, provide a means of producing reasonable estimates of personal wealth utilizing existing data. These estimates are limited by the estate tax filing threshold, which is currently $600,000 in total gross estate. While this threshold is somewhat restrictive, the resulting estimates account for the top 1 to 2 percent of the population. In 1989, these top wealthholders controlled between 25 and 30 percent of personal wealth in the U.S. [Johnson and Schwartz, 1993].

The multiplier is equivalent to a sampling weight, where the probabilities of selection include the probability of being a decedent and that of being included in the SOI sample of estate tax returns. The difficult computation is the probability of being a decedent. Death is not a truly random event and, therefore, the decedent sample is not a simple representative sample of the living population under consideration. The probability that a person will die in a given year depends on many factors. Age and sex have often been taken as the most important factors relating to mortality. However, there is much evidence that the wealthy have mortality rates significantly lower than those of the population as a whole, perhaps due to better access to health care, better nutrition, less hazardous occupations, or better housing [see Menchik, 1991, Kitagawa and Hauser, 1973]. Thus, the probability of being a decedent in our sample has two components, a mortality rate, based on age and sex, and an adjustment, called a differential, which adjusts the mortality rates of the general population for the added longevity of the top wealthholder population. The mathematical expression of the estate multiplier is given in equation (1).

(1) $MULT = 1 \ / \ (p * r * d)$ where:

p = probability of selection to the estate tax sample,

r = mortality rate,

d = rate differential.

There have been several studies as to the best way to compute the differentials. The first researchers to try to make such an adjustment to mortality rates were Daniels and Campion, in preparing wealth estimates for England and Wales in the 1920's [Daniels and Campion, 1936]. Horst Mendershausen was the first U.S. researcher to adjust mortality rates, applying them to IRS data from the 20's and 40's [Mendershausen, 1956]. He used data on the mortality experience of the Metropolitan Life Insurance Company for policies in the $5,000 whole life classification to adjust white, age-specific mortality rates. He was also the first to attempt to isolate insurance values and make an estimate of the cash surrender value, an issue which is discussed in the next section.

Robert Lampman made similar estimates for 1953 based on IRS data [Lampman, 1962]. His estimates carried Mendershausen's work a step further, creating a composite mortality differential adjustment based on three factors: the mortality experience of professional, technical, administrative, and managerial workers for 1950; the 1953 white-male mortality rates; and an average of the 1953 Metropolitan Life data with the data of male Ordinary Life Insurance policy holders. The IRS has been using data from Metropolitan Life comparable to that used by Mendershausen and Lampman to produce estimates since 1962.

Data Sources

There are three main components of the estate multiplier personal wealth estimates: the estate tax sample, the mortality rates, and the rate differentials.

Estate Tax Data

The 1989 SOI Personal Wealth data file is based on Federal estate tax return data compiled by the Statistics of Income Division of the IRS. For the 1989 wealth estimates, the SOI estate tax data were derived from a stratified sample of estate tax returns filed from 1989 through 1991 for individuals who died in 1989. The sample is stratified by year of death, age at death, and by size of total gross estate (TGE). Only estates with a gross estate value of $600,000 or more, the estate tax filing threshold, are included in the sample. All returns filed for both the very wealthy (those with gross estates of $5 million or more) and the young (those under 40 years of age) were selected with certainty.

In the past, wealth estimates, such as the preliminary estimates for 1982, were derived from the SOI sample of estate tax returns filed in a particular year. Because a decedent's estate has up to nine months to file an estate tax return and an extension of six months is not uncommon, returns filed in a given year can include decedents who died in several different years. By sampling returns filed over a 3-year period, nearly all the returns filed for the cohort of 1989 decedents can be represented. We estimate that only 1 percent of 1989 decedents were not included in the study file. These decedents tend to have larger, more complicated estates. We use data for prior years to compute an adjustment to account for these decedents.

The strength of the estate multiplier technique is due, in part, to the nearly complete coverage of the wealthiest portion of the decedent population. Approximately 21,500 estate tax returns for individuals with total assets of $600,000 or more are included in the 1989 SOI Personal Wealth file. Despite the sample size advantages, the limited number of returns filed each year for decedents who were young or very wealthy can make estimates for those sub-groups subject to considerable variance [Smith, 1965].

The number of very young or wealthy decedents tends to vary from year to year and is relatively small in comparison to their representation in the living population. This can result in significant short-term fluctuations in our estimates, attributable solely to the 'sample variance' associated with these two groups. To dampen the effect of these variations, we 'smooth' the sample by including all returns for these individuals filed between 1989-91, without regard to the year of death. The data are then reweighted to represent the true 1989 decedent population.

Three measures of wealth are used in this article: gross estate (or gross assets), total assets, and net worth. The gross estate criterion is a Federal estate tax concept of wealth that does not conform to the usual definitions of wealth. Gross estate reflects the gross value of all assets, including the full face value of life insurance reduced by policy loans but excluding any reduction for other indebtedness. This is the measure used in assessing the estate tax and thus defines the individuals included in the top wealthholder group. The amount of total assets, a lower wealth value, is still essentially a gross measure. Total assets differ from gross assets in that the cash value of life insurance, i.e., the value of insurance immediately before the policyholder's death, replaces the 'at death' value of life insurance included in gross assets [1]. Net worth is the level of wealth after all debts have been removed from total assets.

Mortality Rates

The mortality rates used here are derived from data compiled by the National Center for Health Statistics and are the death rates for white males and females in the United States, by 5-year age groups, for 1989. The rates range from .481 per 100,000 for females, age 20-24 years, to 179.78 per 100,000 for males older than 85. These rates are sample estimates and thus subject to sampling error; however, the sampling error is very small. Thus, we disregard the variance associated with these estimates in our analysis.

Rate Differentials

The rate differentials are derived from information supplied by the Metropolitan Life Insurance Company and are based on the expected vs. realized deaths of their large policy holders. We do not compute sex specific differentials, rather we account only for age at death differences. [2] Our 1989 differential estimates are predicted using differentials in six age categories, for the years 1969, 1971, 1975 and 1978, from Metropolitan Life. We investigated different models for these data, using age and year as the independent variables in an ordinary least squares regression. Year of death was not a significant predictor in any of the models. Our final model is given by equation (2):

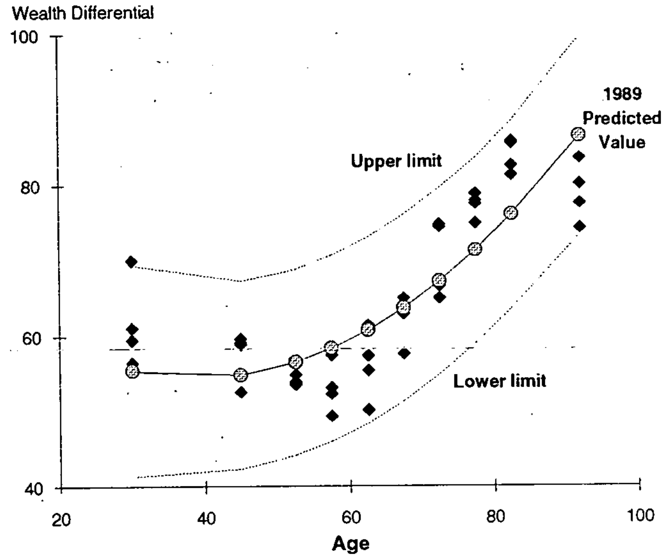
(2) $\text{diff} = 71.808 - 0.89(\text{age}^2) + 0.011(\text{age}) + E{\sim}N(0,5.95).$

The predicted values range from about 58% of the general mortality rate for those under thirty, to 85% for those age 90 and over. The plot in Figure A shows the data, the fitted regression line and the 5% and 95% confidence limits for the individual predictions. We incorporate the inherent model uncertainty in the variance computation, as detailed in the next section.

■ Estimation Methodology

There are two main estimation concerns inherent in sampling applications: the computation of sampling weights and a methodology for computing variance estimates. We first describe the sampling weight computation, which is itself composed of two steps: the computation of the selection probability and adjustments to this probability. As shown in equation (1) and discussed earlier, the selection probability consists of the mortality rate, rate differential, and probability of selection to the SOI sample. We refer to the inverse of this probability as the multiplier. Once all the components

Figure A.--1989 Mortaility Differential Underlying Data and 95% Confidence



are known, we considered adjustments, such as post-stratification and weight trimming.

The extremely skewed distribution of net worth is of particular interest to researchers. Because the underlying sample of estate tax returns was stratified by gross assets, which is not highly correlated with net worth, it would be appropriate to post-stratify. However, the necessary control totals are not readily available. Thus, our strategy was to constrain the tails of the net worth distribution to resemble a Pareto distribution, which is often used in wealth and income models.

For our purposes, the upper tail of the net worth distribution was defined as those individuals with net worth of $250 million or more. In order to determine the parameters of the Pareto, we examined the empirical distribution of net worth implied by the individuals in the *Forbes* 400 for the years 1982-1989. We found that the data approximated a Pareto with parameters varying from 1/2 to 1/3. The SOI data were then divided into the following net worth categories: $250 to $450 million, $450 to $700 million and greater than $700 million.

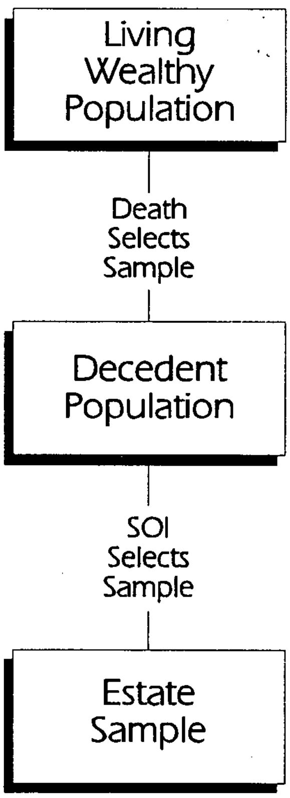
First, we decided to trim the multipliers in the bounded net worth categories at the 3rd quantile. The remaining unbounded category contained an estimated 45 individuals. The multiplier values in this category were fit to a Pareto of parameter 1/3, preserving the final estimate of 45. When these adjustments had been made, the distribution of individuals with net worth of $250 million or more approximated a Pareto, so no further changes were made [3].

Similar adjustments were made for returns with extreme negative net worth (less than -$1 million). These cases were grouped into three categories: -$1 to -$5 million, -$5 to -$15 million, and less than -$15 million. Again, a distribution of the multipliers was computed and the multipliers trimmed at the 3rd quantile in each of the categories.

Variance Estimation

In this section we detail the methodology used in the computation of variance estimates. We quantify the uncertainty inherent in the estate tax data and in the differential used in the computation of the estate multipliers by developing distributional models for these components. These individual measures of uncertainty are used as input to a total error, as described later in this section. Such treatment of the multipliers as an unknown quantity with an error distribution was suggested by Scheuren and McCubbin [1987]. Our overall strategy is to select bootstrap samples that reflect the sampling process from the living population through the selection of the estate sample, incorporating the different estimation steps along the way.

Figure B.--SOI Estate Tax Sample Schematic



The three components of the estate multiplier are the probability of selection to the estate tax sample, the mortality rate, and the rate differential. These approximate a two-stage sampling scheme, where the first stage is death, with the probability of selection equal to the mortality rate times the differential. The second stage is the SOI sample selection (see Figure B). Note that for both stages, the sampling unit is an individual. We have nearly complete knowledge of the probability of selection to the estate tax sample. The sample is a stratified random sample with sample rates varying from 7 to 100 percent. We also have some limited information for the entire population of estate tax filers. It is worth pointing out that the sample is post-stratified to the population of filers. This allows us to adjust for returns that were misclassified due to keying errors prior to sample selection. Still, there is sampling error associated with the SOI sample.

Less is known about the probability that an individual will die in a particular year. While we have mortality rates, a degree of uncertainty must be attributed to this process, as well. (We have already detailed our efforts to 'smooth' the variance of the sample selection of the young and wealthy in the previous section.) Also, because the differentials are modeled using a time series of data, they add to the uncertainty of the multiplier [4]. In order to measure the contribution of each of these components of uncertainty (sampling and modeling), we choose the bootstrap method of variance estimation.

The bootstrap method of variance estimation is a resampling technique where many random samples are drawn, with replacement, from the original sample. The bootstrap samples are selected using the same sampling scheme as was used for the original sample. These bootstrap samples are used to produce estimates which are then combined to produce estimates of bias and variance for desired statistics. The advantage of a resampling method is that it is possible to estimate the variance of both linear and nonlinear statistics and to incorporate uncertainty due to the estimation process. Additionally, inferential analysis, such as the computation of confidence intervals, is not limited to the assumption of an underlying normal distribution; the actual distribution of the bootstrap sample estimates can be studied.

In order to select bootstrap samples, we must recall that the estate tax sample is derived from a two-stage selection process. This sample is not a traditional cluster sample, because the sampling unit is the same for both stages and the samples are selected independently of each other. However, the estate sample is a subsample of death's decedent sample. We assume death's sample is stratified by age at death and sex. In order to capture the sampling variance due to death's selection, we recreated death's decedent sample. Based on the SOI sample weights, records were duplicated, creating a decedent sample of 53,000 (the total number of estate tax filers who died in 1989). We then resampled, with replacement, in the age/sex categories we used in assigning mortality rates--our assumed stratification of death's sample.

At this point, we have a 'bootstrap decedent sample' reflecting death's selection. From here, we reselect an SOI estate tax sample, according to the original SOI sample design. We now have a bootstrap estate sample incorporating both the sampling error attributed to death, and that attributed to the SOI sample. Next, we choose a differential from the estimated distribution, incorporating the model error. We constrain the differential choice to be within the computed 95% confidence interval, seen previously, to keep the resulting values in a feasible range. Finally, using the known probability of selection to the estate sample, the assumed mortality rate, and the estimated differential, we compute the multipliers. We bound the multipliers for decedents with net worth over $250 million or net worth less than -$1 million, as discussed earlier. These adjustments are data driven, the third quantile bound was computed for each bootstrap sample.

Figure C shows a box plot of 11 bootstrap estimates for total assets. The median of the bootstrap estimates is $5.37 trillion, with a high of $5.50 trillion and a low of $5.25 trillion. As expected, the distribution of the estimates is slightly skewed. Because there are no control totals to use in computing the multipliers, the estimate of the number of individuals is also of interest. Figure D is a box plot of 11 bootstrap estimates of the number of individuals with total assets over $600,000. The median of the estimates is 3.06 million, with a high of 3.08 million and a low of 3.04 million. The distribution is much tighter than that of the sum of total assets.

In computing the estimates, it was possible to measure the variance associated with each of the multiplier components. The post-stratification adjustments, which included the weight trimming, decreased the variance of the sum of total assets by 54%. The resulting estimate was about 1.4 percent less than the original, unpost-stratified estimate. The variance of the frequency estimate decreased 16% as a result of post-stratification; the effect on the value of the estimate was negligible.

Allowing the mortality rate differentials to vary within the 95 percent confidence interval suggested by the model increased the variance of the total asset estimate by about 29 percent; the variance of the frequency estimate increased by less than 2 percent. In both cases the resulting estimates were almost 4 percent higher than the estimates based on a fixed differential. Further investigation revealed that calculating the differentials based on the age of each decedent using equation (2), rather than assigning them in broader age categories, as was done

Figure C.--Box Plot of Aggregate Total Assets Using 11 Bootstrap Estimates

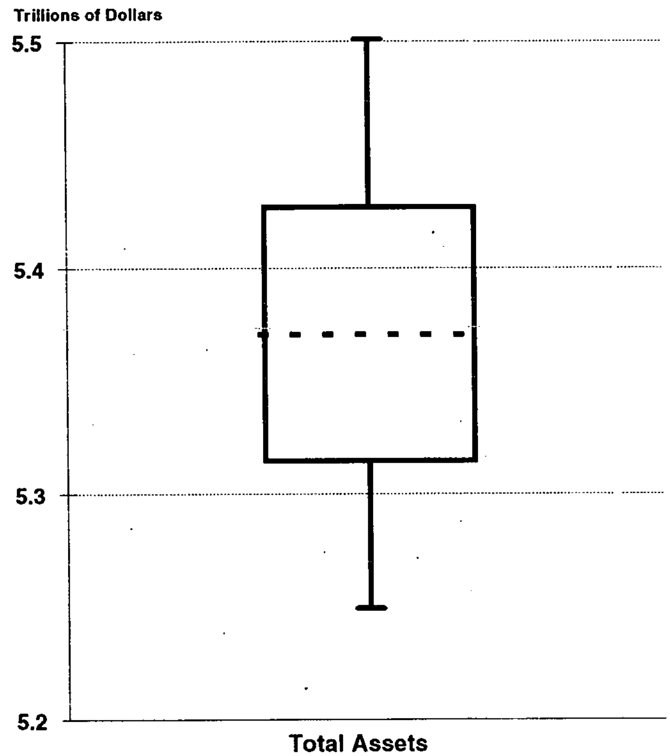
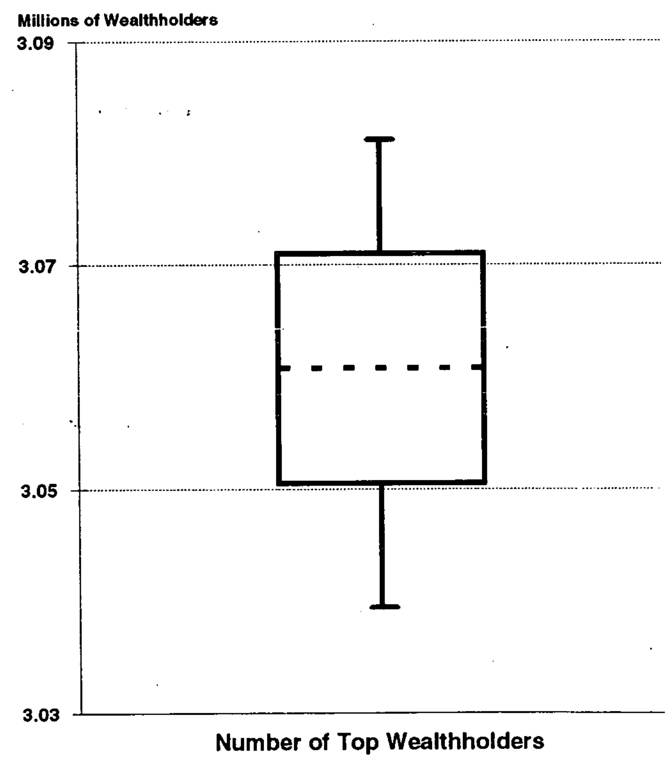


Figure D.--Box Plot of Number of Top Wealthholders using 11 Bootstrap Estimates



previously, further increases the estimate of both the number of top wealthholders and the total value of their assets.

■ Comparison to Other Data

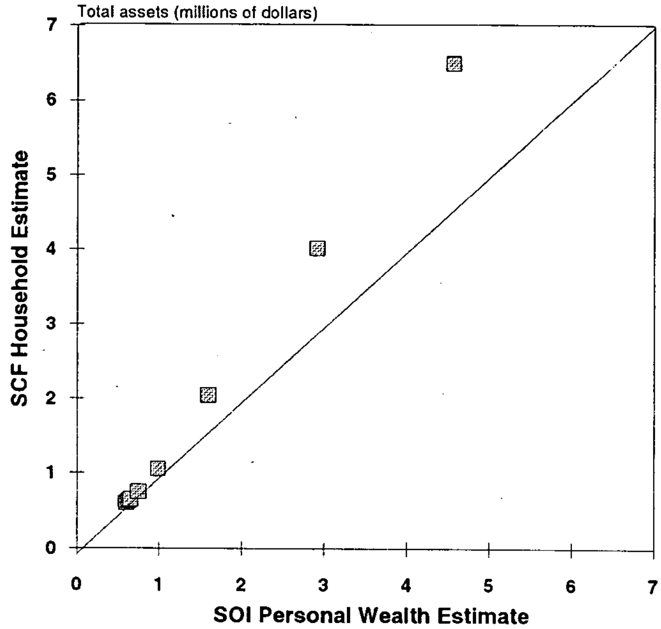
Having created estimates of individual or personal wealth, the next step is to validate them, using an independent data source. As mentioned, there is very little information on the wealth of individuals in the U.S., particularly for those in the upper end of the wealth distribution. One excellent source of wealth data, however, is the Survey of Consumer Finances, sponsored triennially by the Board of Governors of the Federal Reserve System.

The Survey of Consumer Finances (SCF) is a household survey from which estimates of wealth, income, savings, etc. for the entire nation can be derived. One main objective of the SCF is to provide a good representation of the entire wealth distribution. In order to do this, the SCF incorporates a dual frame sample. One sample is a multi-stage area probability sample; the other is a list sample stratified by a measure of wealth and sampled disproportionately [see Kennickell and Woodburn, 1992]. It is particularly appropriate to use estimates derived from the SCF to validate the SOI estimates because of the supplemental coverage of wealthy individuals provided by the list sample [5].

Because the SCF produces household estimates of wealth, while the estate multiplier technique produces estimates of individual wealth, it is not possible to compare aggregate totals directly [see Scheuren and McCubbin, 1987]. However, there should be some relationship between the distributions of the estimated populations. This comparison can be shown graphically using quantile-quantile (QQ) charts, which compare the cumulative percentage of individuals over a fixed set of percentiles [see Wilk and Gnanadesikan, 1968, for an explanation of this technique]. If the distributions of these functions are exactly alike, the plot will be a straight line, passing through the origin, with slope=1. If the variances are the same, the y-intercept represents the difference in the means. The slope represents the ratio of the variances. If the distributional forms are dissimilar, the plot will be nonlinear.

Figure E is a QQ comparison for estimates of total assets for individuals (SOI) vs households (SCF) with at least $600,000 in total assets. The linear arrangement of the points indicates that the two distributions are of

Figure E.--QQ Comparison, SCF vs. SOI Estimates for Households/Individuals with Total Assets of $600,000 or More



similar functional form. The slope is greater than one, which shows that the values of the SCF estimates rise more quickly than those derived from the estate data.

The linear relationship between the two sets of estimates observed in the QQ plot suggests that the underlying distributions are similar, albeit, derived in different ways. It is possible to carry the comparison further by using additional SOI data to create a household data base from the individual wealth file for households with total assets of at least $600,000. We start with the assumption that single individuals, whether they are widowed, separated, divorced or never married, each represent a household. We are left with devising a way to create 'families' from the data for married individuals.

The household wealth of a married couple can be separated into three parts: assets belonging solely to one of the two spouses (WFEM or WMALE) and assets held jointly (WJOINT). The data in the SOI estimates represent individuals for which WFEM + 1/2 WJOINT (for females) or WMALE + 1/2 WJOINT (for males) was at least $600,000. The value of the second spouse's assets (WMALE for females or WFEM for males) is missing and must therefore be imputed.

We first took up the task of imputing WFEM for the males in our file, simply because there were more records for married males in the database than married females.

In order to impute a value for WFEM, we made the following assumptions:

1. Some, but not all, of the married individuals in the SOI estimates are married to each other.
2. The separate assets of the married females in our file were representative of those belonging to the spouses (WFEM) of the males in the file.

In general, imputation of missing values is most effective when based on a model derived from a distribution of known values [Little, 1988]. Since WFEM is missing for all cases, we have no information on how WFEM relates to WMALE or, for that matter, to any other data for a given male. Therefore, an explicit model was not feasible. We instead chose the hotdeck procedure within adjustment cells [Hinkins and Scheuren, 1986]. Records for married males for which a value of WFEM was to be imputed were matched to records for females (donor records) in the same adjustment cell. The missing value was estimated using the known value from the donor records.

In order to implement this procedure, the donor records must first be divided into cells. The original SOI sample Total Gross Estate (TGE) categories were used, creating 3 strata: TGE under $1 million, $1 million under $5 million and $5 million or more. Each of these strata was further divided into four quantiles based on the distribution of joint assets within that stratum, creating a total of 12 cells. Records for the males were divided into the same cell categories. A value of WFEM was then chosen randomly, with replacement, from a donor record, for each married male in a corresponding cell. Repeated applications of the procedure showed that the variance attributable to the imputation process had a relatively minor effect on the distribution of the value of Total Assets in the resulting data set. This variance is incorporated later.

The results of the hotdeck procedure produced a data set containing estimates of households where WMALE + 1/2 WJOINT was at least $600,000. While we assume that some of the males and females in our file are members of the same household, there remain a number of households for which the female spouse's assets (WFEM + 1/2 WJOINT) totaled at least $600,000, but the husband's did not. We assumed that households in which females owned separate assets of at least $600,000 (WFEM $\geq$ $600,000) best represented these missing families. The males on our file which best represented their spouses were those for whom WMALE < $600,000. A weighting adjustment was made to account for these additional 'families.'

This final adjustment gave us a file of households for which at least one individual owned $600,000 or more in gross assets. We were not able to represent married households for which each individual owned less than $600,000 but where the couples' combined assets totaled $600,000 or more. These households are, however, included in the SCF estimates and, thus, direct comparisons of the SOI and SCF households at the $600,000 threshold are not meaningful.

The effect of these 'missing' families on the frequency and dollar estimates should diminish at higher total asset thresholds; the SCF and SOI estimates should eventually converge. Figure F gives frequency and dollar estimates of total assets for different thresholds between $600,000 and $1 million. The frequencies converge between $850,000 and $900,000. At this level their aggregate estimates of total assets differ by about $890 billion, with

Figure F.--SCF vs SOI Household Estimates of Aggregate Total Assets for Increasing Total Asset Thresholds

Total asset threshold	SCF estimate		SOI estimate		Percent difference between SCF and SOI	
	Number	Amount	Number	Amount	Number	Amount
$600,000	3.93	8.89	2.94	7.26	25.19 %	18.34 %
$650,000	3.51	8.62	2.80	7.20	20.23	16.47
$700,000	3.10	8.36	2.67	7.11	13.87	14.95
$750,000	2.93	8.24	2.58	7.04	11.95	14.56
$800,000	2.61	7.99	2.48	6.97	4.98	12.77
$850,000	2.42	7.83	2.39	6.90	1.24	11.88
$900,000	2.30	7.71	2.32	6.83	-0.87	11.41
$950,000	2.16	7.59	2.25	6.76	-4.17	10.94
$1,000,000	2.03	7.46	2.17	6.69	-6.90	10.32

the SOI estimate about 12 percent less than that of the SCF. A QQ plot of total assets for households with at least $900,000 in assets is given in Figure G. The distributions are more similar than before, but the distribution of the SCF estimates still increases more rapidly than that of the more conservative SOI estimates. Interestingly, some researchers have estimated that the pre-audit values of some estates increase by as much as 10 percent after an audit has been completed [Harris, 1949], while a more recent study found that the increase was between 2 and 4 percent, overall [McCubbin, 1987]. It is also reasonable to expect that values which are acceptable for administrative purposes may be more conservative than those given during a survey interview.

Finally, by applying the variance methodology previously described, including the additional variance due to the imputation of "families," we can see that much of the difference in the two estimates is attributable to the uncertainty associated with the two estimation techniques. Figure H shows a box plot of 11 bootstrap estimates of the value of assets for households at or above the $900,000 threshold. The overall variance of the SCF estimate is much greater than that of the SOI estimate, largely because of the much smaller sample size on which the SCF estimates are based. (Remember that at this threshold, the frequency estimates are nearly the same.) This graph shows that the more conservative SOI estimate of the value of total assets is within the survey's margin of error.

Figure G.--QQ Comparison, SCF vs. SOI Estimates for Households with Total Assets of $900,000 or More

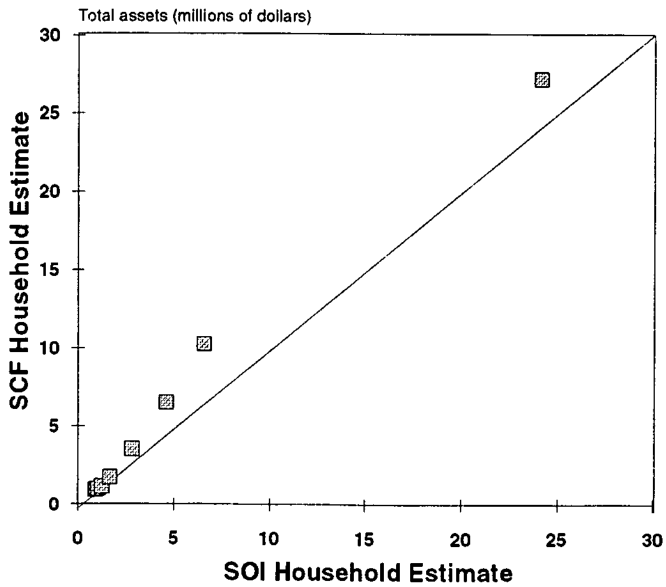
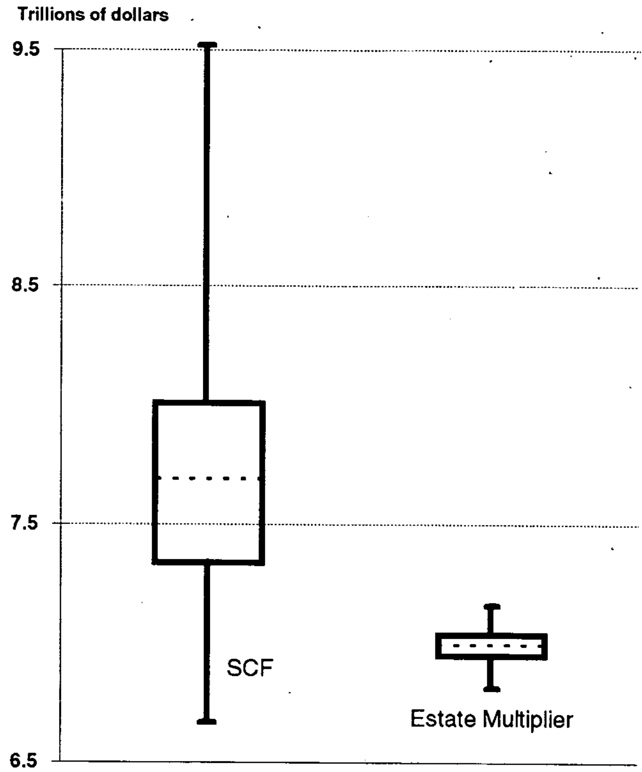


Figure H.--Comparison of Bootstrap Estimates: Aggregate Total Assets for Households with Total Assets of $900,000 or more



■ Summary and Future Research

In his 1965 paper presented to the American Statistical Association, James Smith suggested several areas for future studies relating to the estate multiplier technique and wealth estimation. Likewise, Scheuren and McCubbin gave a similar set of suggestions in their 1987 paper. In closing, it seems appropriate to re-visit their "wish lists" and give a progress report on those things we have not already specifically addressed.

1. The major information gaps in the Federal estate tax return data had been missing ages and the absence of a cash surrender value of life insurance. The first problem, that of missing ages has been virtually eliminated through the use of death certificate data, now required of each filer, to supplement data reported on the return itself. For our present sample, there were a mere 22 decedents for whom ages were imputed.

   Estimating the cash value of life insurance remains a problem which will only get worse, in light of the many new products being marketed by the industry.

Some of these products are sophisticated investment instruments, the return on which is determined by an individual's attitude toward risk as well as by market performance. We are continuing to investigate ways to capture more information from the supplemental data filed with each tax return in order to improve in this area.

2. Data captured from estate tax returns is prior to an audit. C. Lowell Harriss estimated that the value of some estates may increase by as much as 10% after audit. A more recent preliminary study by Scheuren and McCubbin suggest that the difference may be much less significant. SOI is planning a more extensive study, to begin in 1994, which will look at this important issue.

3. What is the income of top wealthholders and their heirs? What is the relationship between *inter vivos* giving and the transmission of assets at death? These issues are very important to the estimation of wealth and are being addressed through a series of estate collation studies, beginning with a sample of 1976 decedents. These studies link income, gift and fiduciary tax returns filed for decedents and their heirs, for several years prior to a decedent's death, and, in the case of heirs, several years hence. We are currently beginning to process data for 1989 decedents and are planning a larger scale study for 1992 decedents. The advent of the SOI individual income tax panel sample (see Czajka and Walker, 1990) will also offer important opportunities to follow income patterns of individuals over long periods of time and then relate those data to their estate tax returns at their death.

4. Finally, of course more research is needed into the appropriate mortality rates. We have shown that allowing the mortality rate differentials to vary within age categories can have a significant influence on the final estimate and its variance. Further, the practice of assigning the differentials within broad age categories seems to bias the resulting estimates downward. We are looking at several sources for more detailed information on the influence of wealth on mortality, the most promising being the National Longitudinal Mortality Study sponsored by the National Institutes of Health.

Notes

[1] The cash value of life insurance included in total assets and net worth was approximated based on the face value of life insurance reported on the estate tax return and on the decedent's age. This was done by applying an equity valuation ratio in the form:

$$\frac{\text{Cash Value of Life Insurance}}{\text{Face Value of Life Insurance.}}$$

This ratio was developed based on two independent sources of data. The first was a study conducted by the Institute for Life Insurance, which looked at the life insurance policies which had been reported in the estates of Federal estate tax filers in mid-1971. The second was the Federal Reserve Board's 1989 Survey of Consumer Finances. This survey asked respondents to approximate the total cash and surrender values of their life insurance values. Only households with total assets of at least $600,000 were considered.

The results generated from each of these sources were encouragingly similar. A simple regression in which age was the independent variable was used to predict the values used in our estimates. No attempt was made to adjust for the presence of term insurance in our data. The same set of ratios was used for both males and females, due to lack of sex-specific data.

Life Insurance Equity Values

Age	Equity Ratio
Under 40	3.9%
40 under 45	8.9
45 under 50	14.0
50 under 55	18.4
55 under 60	24.2
60 under 65	31.1
65 under 70	38.6
70 under 75	47.0
75 under 80	56.1
80 and over	82.5.

[2] We have derived a single set of mortality rate differentials, based on the decedent's age, which were used for both males and females. Based on the knowledge that these groups have very different mortality rates, this may be inappropriate. A simple test of this is to compare the estimates of community property for married males and females living in states where such property is common [Scheuren,

1975]. It seems reasonable to expect that the frequency and dollar estimates of community property would be roughly equal between men and women if the differentials have been assigned appropriately.

The table below shows estimates for all community property states combined and for California, the state for which we had the largest sample. In both cases, the aggregate dollar estimate is higher for women than for men; the frequency estimates are reasonably close. Based on these results, no further adjustments were made, although more research is needed in determining the characteristics of the mortality differential between the wealthy and the general population.

Community Property Estimates

(Amounts are in billions)

	Males		Females	
	Number	Amount	Number	Amount
All	270,000	$741	266,000	$949
Calif	157,000	$472	158,000	$584

[3] It is interesting to note that the adjusted data base estimated that there were nearly 400 individuals with net worth greater than $250 million, the *Forbes* 400 cut-off in 1989.

[4] Additionally, there is a degree of uncertainty in the estimation of the various components of wealth (real estate, stock, bonds, etc.) due to errors introduced during data capture. These errors are, however, beyond the scope of this paper.

[5] The Survey of Consumer Finances does not include any individuals with net worth greater than $250 million. We therefore constrain the SOI estimates in the following comparison to the same upper bound.

References

Avery, Robert B. and Kennickell, Arthur B. (1992), "Household Saving in the U. S.," *Review of Income and Wealth.*

Czajka, John L. and Walker, Bonnye (1990) "Combining Panel and Cross-Sectional Selection in an Annual Sample of Tax Returns," *Statistics of Income and Related Administrative Record Research: 1988-1989*, pp. 127-132.

Daniels, G.W. and Campion, H. (1936), *The Distribution of National Capital*, Manchester: Manchester University Press.

Greenwood, Daphne, (1981), "A Method of Estimating the Distribution of Wealth Among American Families," *1981 American Statistical Association Proceedings, Section on Survey Research Methods*, pp. 472-475.

Harriss, C. Lowell (1949), "Wealth Estimates as Affected by Audit of Estate Tax Returns," *National Tax Journal*, December 1949.

Hinkins, Susan, and Scheuren, Frederick (1986), "Hot Deck Imputation Procedure Applied to a Double Sampling Design," *Survey Methodology*, Volume 12, pp. 181-196.

Internal Revenue Service (1976), *Statistics of Income—1962, Personal Wealth Estimated from Estate Tax Returns*, U.S. Department of the Treasury.

Johnson, Barry W. and Schwartz, Marvin (1993), "Personal Wealth, 1989," *Statistics of Income Bulletin, Spring 1993*, Volume 12, Number 4.

Kennickell, Arthur 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, pp. 1-18.

Kennickell, Arthur B. and Woodburn R. Louise, (1992), "Methodological Issues in the Estimation of the Distribution of Household Net Worth: Results From the 1989 Survey of Consumer Finances," Federal Reserve Board, working paper.

Kitagawa, Evelyn M., and Hauser, Philip M. (1973), *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*, Cambridge: Harvard University Press.

Lampman, Robert J. (1962), *The Share of Top Wealthholders in National Wealth, 1922-56*, Princeton: Princeton University Press.

Little, R.J. (1988), "Missing-Data Adjustments in Large Surveys," *Journal of Business and Economic Statistics*, 6, Number 3: 287-301.

McCubbin, Janet (1987), "Improving Wealth Estimates Derived from Estate Tax Data," *Statistics of Income* *and Related Administrative Record Research: 1986-1987*, U.S. Department of the Treasury, Internal Revenue Service.

Menchik, Paul L. (1991), "Economic Status as a Determinant of Mortality Among Nonwhite and White Older Males: or, Does Poverty Kill?" Institute for Research on Poverty, Discussion Paper # 936-91.

Mendershausen, Horst (1956), "The Pattern of Estate Tax Wealth," in Raymond W. Goldsmith, Dorothy S. Brady, and Horst Mendershausen (Eds.), *A Study of Savings in the United States*, Vol. III, Princeton: Princeton University Press.

Scheuren, Fritz (1975), "Historical Perspectives on IRS Wealth Estimates With a View to Improvements," Remarks delivered at the Workshop on Estate Multiplier Estimates of Wealth, sponsored by the National Bureau of Economic Research.

Scheuren, Fritz, and McCubbin, Janet (1987), "Piecing Together Personal Wealth Distributions," *Statistics of Income and Related Administrative Record Research: 1986-1987*, U.S. Department of the Treasury, Internal Revenue Service.

Smith, James (1965), "Estimating the Wealth of Top Wealthholders from Estate Tax Returns," *1965 American Statistical Association Proceedings, Section on Business and Economics Statistics*.

U.S. Bureau of the Census, *State Population and Household Estimates: July 1, 1989,* Current Population Reports, Series P-25, Number 1058, .

Wilk, M.B. and Gnanadesikan, R. (1986), "Probability Plotting Methods for the Analysis of Data," *Biometrika*, Vol. 55, pp. 1-17. (As described in Hoaglin, David C. (1985), "Using Quantiles to Study Shape," *Exploring Data Tables, Trends and Shapes*, New York: John Wiley and Sons, pp. 417-460.) ■