SOME PRELIMINARY RESULTS FROM A VALIDATION STUDY OF THE ESTATE MULTIPLIER PROCEDURE

H. Lock Oh and Fritz Scheuren*
Social Security Administration

The estate multiplier procedure has long been employed by economists and statisticians to estimate personal wealth [e.g.,1]. While there is a widespread recognition of the fact that the technique has potentially serious flaws, repeated calls for a vigorous evaluation of the method [e.g.,2] have generally been unavailing. A number of projects, however, are now underway which offer some promise in this regard [3,4]. The present paper is a report on one such effort that was recently undertaken using data from Social Security's Continuous Work History Sample.

ESTATE MULTIPLIER PROCEDURE

Before discussing this validation study and the results that were obtained, it might be useful to review some of the salient features of the multiplier procedure. The estate multiplier technique employs information from the records of decedents (usually financial records) to describe the characteristics of living individuals. To construct estimates, for each decedent, the data are weighted by the inverse of the mortality rate thought to be appropriate for the demographic category to which the individual belonged.

Illustration.--Consider the estimation of average house value for white males, 35 to 44 years of age, living in the State of Massachusetts in 1973. Assume we have at hand data on the number of 1973 decedents in this category, their house values, as obtained from some source (such as probate records) and their respective mortality rates. The first three rows of figure 1 provide some hypothetical data on these decedents. In particular, for the 35 to 39 year age group, there are 80 decedents, and their average house value is $40,000. For the 40 to 44 year age group, there are 100 decedents who had total house values of $4.5 million, or $45,000, on the average.

The next two rows deal with mortality experience. By definition, the mortality rate for a given demographic category is the fraction of the persons in that category who died. If one assumes that the occurrence of death for any individual in such a group is equal to that for any other member, then, the mortality rate is equivalent to the sampling fraction in a stratum of a stratified survey design. This analogy to probability sampling is generally used [5,6] to "justify" the estimation procedure employed in the estate multiplier technique. In any case, given the mortality experience shown in figure 1, the estate multiplier procedure would weight the information for the 35 to 39 year-olds twice as heavily as for those 40 to 44. The balance of the illustration shows how this is done.

Advantages and Disadvantages of Method.--There are several advantages that the multiplier enjoys which make it an attractive alternative to probability samples employing household survey techniques. First, the information is generally obtained from records that have been very carefully prepared. Second, if the records are for estate tax or probate purposes, legal sanctions exist which would tend to further reduce misreporting problems. Coverage errors and errors arising from nonresponse are also lessened considerably because of the routine compliance procedures associated with the administration of the law. Questions of ownership and valuation do arise, of course; but, on the whole, at least in the case of the U.S. Federal estate tax returns, content errors are believed to be quite small [7].

The main disadvantage of the technique is that the estimates are not based on probability samples. In particular, the "randomization" is not under the control of the analyst, and, hence, he must guess about its nature. Thus, a subjective element is introduced which might have a crucial impact on the results.

Figure 1.--Illustration of the Estate Multiplier Calculations

Item	Age (in years)		
	35-39	40-44	35-44
DECEDENT DATA			
Decedent Sample Size....	80	100	180
Total House Value (in millions of dollars)....	3.2	4.5	7.7
Average House Value (in dollars).....	40,000	45,000	42,728
MORTALITY EXPERIENCE			
Mortality Rate.....	2/1000	4/1000	-
Estate Multiplier.....	500	250	-
ESTIMATES USING ESTATE MULTIPLIER			
Total Population Count..	40,000 (80x500)	25,000 (100x250)	65,000
Estimated Total House Value (in millions of dollars).....	1,600	1,125	2,725
Estimated Average House Value (in dollars).....	40,000	45,000	41,923

As usually carried out, the multiplier technique stands, or falls, depending on whether or not the following assumptions hold:

1. the characteristics to be measured have not been distorted by the sampling process (i.e., by the occurrence of death);
2. the average mortality experience of the population about which inferences are being made has been adequately accounted for in the rates being used; and, finally,
3. the extent to which an individual's probability of death differs from the average for his demographic group is not related to the information one wishes to estimate.

Of course, it has generally not been possible to adequately check these assumptions. Even in the validation study we are about to describe, we have only been able to isolate the net effect of failures in these assumptions.

DESCRIPTION OF VALIDATION STUDY

One way to validate the multiplier procedure is to compare it with another measurement technique in which we have more (or complete) faith. This turns out to be a very formidable problem. In the context in which the multiplier procedures are usually conducted, only very limited success has been achieved so far.

It is possible, however, to examine the multiplier method in situations where it would not normally be used because other estimators are available. While not quite in the needed context, such situations do afford us a test of the procedure. Of course, the nature of the situation one studies naturally limits one's inferences about the validity of the procedure, but this should (and did) not deter us.

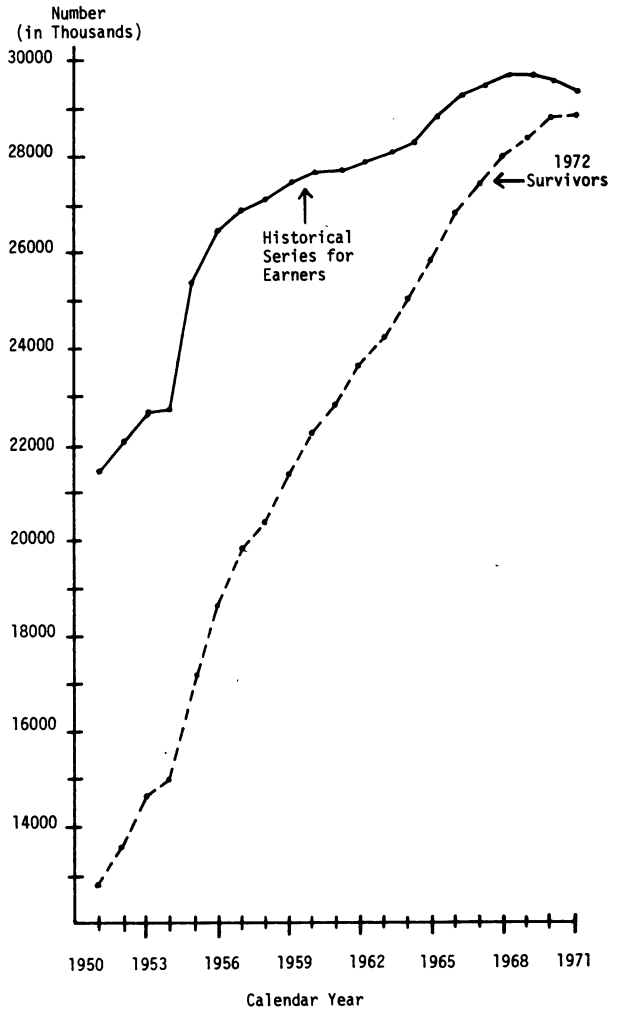
Use of Social Security (SSA) Records for Validation Purposes.--SSA maintains longitudinal records on each worker's social security covered taxable earnings up to the taxable maximum for any given year. Annual summaries from these administrative files are published in the Statistical Supplement to the Social Security Bulletin [8].

The basic study design called for a comparison between this overall historical information for the period 1951-1971 and estimates obtained by tabulating an appropriately weighted sample of persons in the Social Security system who were identified as having died in 1972. 1/ If the multiplier procedure were valid, it should be possible to estimate the earnings distribution in some prior year (say 1960), by employing the earnings histories obtained from the longitudinal files for 1972 decedents. However, a difficulty arises which must be faced. It has to do with the fact that the 1972 sample could only be expected to estimate the earnings distributions of individuals who survived to 1972. In particular, individuals who died prior to 1972, but who had earnings in, say, 1960, cannot be estimated using just the 1972 decedents.

Survivor Estimates.--Now, if all deaths for all persons in social security covered employment were always reported to SSA, the published historical series could be adjusted directly, so that only earners who had survived to 1972 would be included. Unfortunately, there is no necessity for all deaths to be reported. In many cases, there is an economic incentive to notify Social Security (if the individual has worked long enough to be eligible for a lump sum death benefit), but, even then, the death is not always reported.

There were two consequences of this: one is that, although the reporting has improved over the years, the 1972 decedent sample's coverage was such that we could only look at a rather restricted universe--males 35 years or older in 1972. Even with this restriction, there was still some coverage error, with only about 95 percent of all 1972 deaths among males aged 35 or more being

Figure 2. --Number of Male SSA Earners, 35 Years or Older: Comparison Between Historical Series and 1972 Survivors, 1951-1971



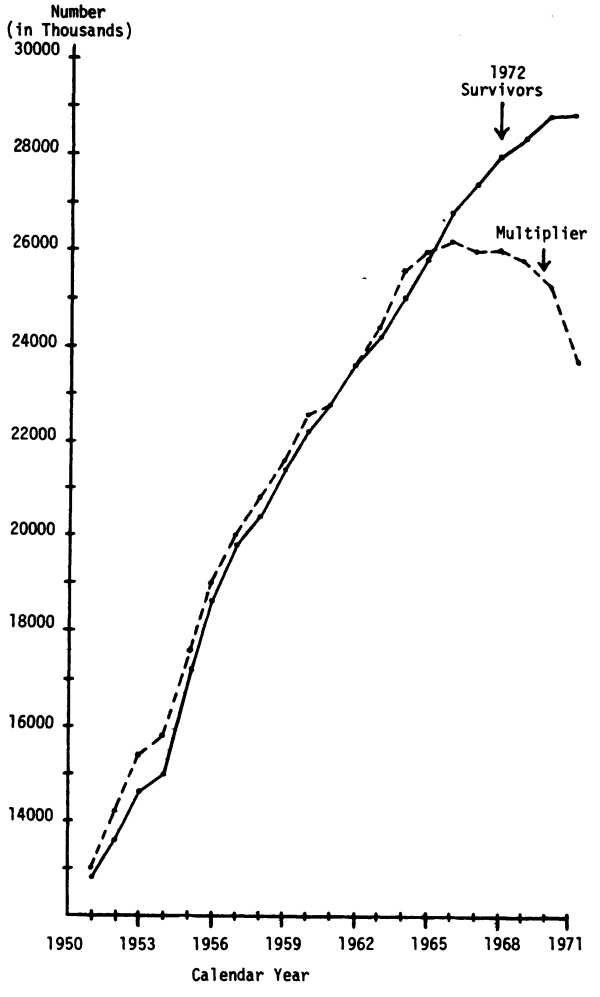
reported to SSA. Five year age-race specific coverage adjustments could be made for 1972 decedents;2/ however, it was not possible to use a direct method of calculating 1972 survivors, since the coverage error not only was greater in earlier years, but it is cumulative.

To "get around" this problem, we had to synthetically derive cohort survival rates by single year of SSA age (race and sex) for the period 1951 through 1971. The details of how these survival rates were calculated are provided in [9], which is available on request.

Figure 2 compares the 1972 survivor counts obtained by this process to the corresponding distribution of the number of persons in SSA's historical series for the years 1951 through 1971. There are no surprises here. For example, of the 21.3 million male earners 35 or older in 1951, the chart shows that 12.7 million, or only about 60 percent, survived to 1972. On the other hand, as might also be expected, the fraction of 1969 earners who survived to 1972 was over 96 percent.

The 1972 survivor totals are, of course, subject to various errors. On the whole, though, they probably afford a reasonable standard for the initial check of the multiplier procedure given in the next section.

Figure 3. --Number of Male SSA Earners, 35 Years or Older: Comparisons Between 1972 Survivors and Multiplier Estimates, 1951-1971



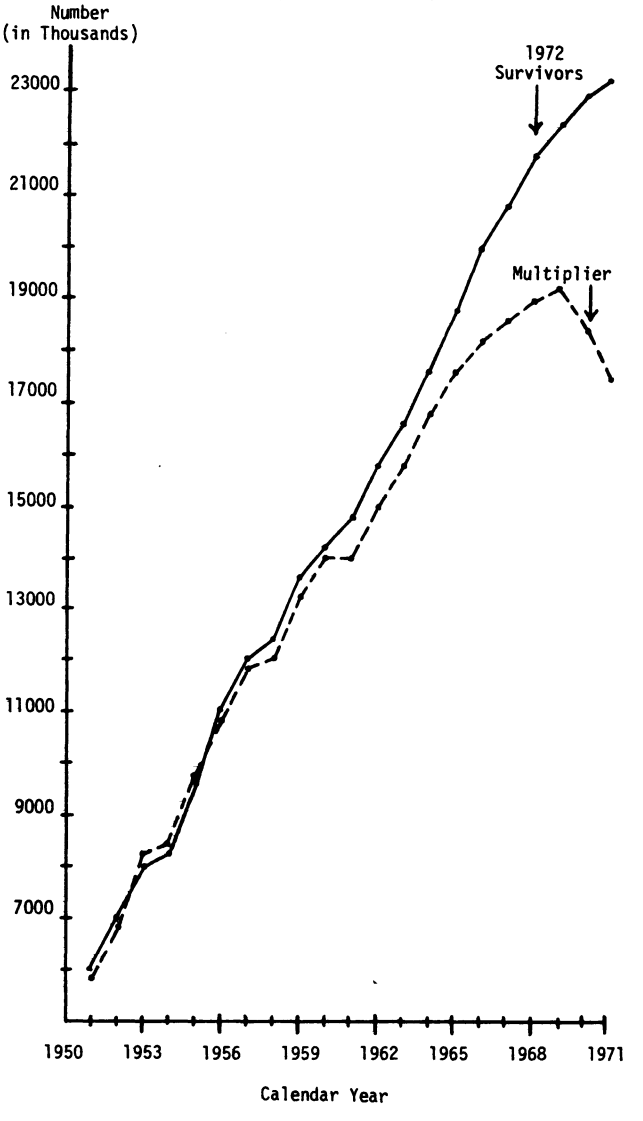
INITIAL RESULTS OF STUDY

For the comparisons we will be making in this paper, the demographic categories in which the multiplier estimates (and the survivor figures) were calculated include only age, race, and sex. No adjustment for differentials by earnings class (or, indeed, between earners and nonearners) has been made. This limitation should be kept in mind as we look at figures 3 and 4 below.

Overall 1972 Survivor Series and Multiplier Estimates.--Figure 3 compares the number of survivors shown in the previous chart to multiplier estimates based on decedents' records for 1972.

Notice that the multiplier estimates are generally quite similar to the survivor totals. For the

Figure 4. --Number of Male SSA Earners, 35 Years or Older with Earnings of $3600 or More: Comparison Between 1972 Survivors and Multiplier Estimates, 1951-1971



period before 1966, the two graphs track each other very closely. In fact, for each year, they are within sampling error.3/ However, the multiplier procedure fails to provide reasonable estimates after 1967. Apparently, as early as six years before death, the onset of health problems causes a decrease in earnings. We expected this behavior for 1971 and, to a lesser extent, for 1970. It was somewhat surprising to us, though, to see that the phenomenon begins so long before death and that its effect is so marked.

1972 Survivor Series and Multiplier Estimates for "High" Earners.--Figure 4 examines the multiplier estimates for persons with earnings of $3,600 or more.4/ Here, the multiplier appears to do well only until about 1960, while in figure 3, the estimates were in rough agreement as late as 1966. We speculate that not only are health problems a factor, but that there may also be a favorable differential at work in the mortality of these "high" earners relative to all social security account number holders. In this connection, it should be pointed out that, if there were such a differential, then, the survivor totals being used as a standard would be too low; hence, the gap shown in figure 4 would be an understatement.

FURTHER RESULTS AND SOME CONCLUSIONS

While this presentation has probably not emphasized it sufficiently, there are a number of "methods" issues imbedded in the approach we have taken. It is for this reason that we have labelled the findings as preliminary. Nonetheless, we do think some generalization may be warranted. Clearly, the estate multiplier procedure cannot be said to have been "validated" by the tests offered in this paper. However, proponents of the technique should not despair. For one thing, the impact of "health problems" would probably not be as severe for wealth variables as they are for earnings. On the other hand, the potential effect of differentials in mortality rates could be much greater in estimating wealth.

This paper has only touched on some of the overall results in the complete study. An extensive appendix is available to anyone interested in pursuing the matter further.

AN AFTER WORD

When we prepared this paper for the Proceedings, it struck us that we had not adequately stressed that, in our opinion, the multiplier technique has been overused. In particular, we feel the technique is unsafe to employ alone or without introducing external (and internal) checks. (See [4].) It is one thing to examine the method as an intellectual curiosity (as was done in this paper). It is entirely another matter to rely on it in situations where important decisions have to be made.

FOOTNOTES

* The authors would like to thank Wendy Alvey, Faye Aziz, Beth Kilss, and Bob Yuskavage for the computational and other assistance they provided. Helpful editorial comments were given by Keith Gilmour. The text was typed by Catherine Murphy.

1/ The decedent sample used was the 1% Continuous Work History Sample (CWHS) for the period 1937-1972. Information was also tabulated on all earners for each year from 1937 to 1971 using the 1937-71 0.1% CWHS. These tables were needed to augment the published historical data. (A full set of the tabulations will be made available upon request.)

2/ The assumption was made that, for a given race-age group, the earnings histories of covered and uncovered 1972 decedents were roughly the same. This assumption is undoubtedly false, but, because the uncovered group is so small, the differences which exist are not believed to have materially affected the overall outcomes of the present test. A study matching death certificates to social security records is planned to examine the characteristics of the uncovered group.

3/ For the "survivor" totals, the variance was fairly easy to obtain, since, at least at the national level, for the statistics examined, the CWHS (upon which the figures were based) can be treated as roughly a simple random sample. (See [10], table 1.) To estimate variances for the multiplier totals, we had to resort to a super-population model. Under this model we assumed that:

1. each calendar month's deaths were drawn independently of every other month's deaths;

2. the monthly samples can be divided into six strata--

   A. January and December
   B. February and March
   C. April and November
   D. May and October
   E. June and July
   F. August and September; and

3. within each of these 6 strata, the probability of an individual's death during a month is equal to some constant which is the same for both months. [12]

One advantage of this formulation is that it does not introduce any assumption about the nature of the bias in our estimator. It also takes some account of the seasonality which exists in the death rates. The approach is a "good bit" better than simply assuming that we are engaged in stratified sampling where the strata are race-age groups. However, it

must be admitted, we are somewhat dissatisfied with this method, and, in the future work, we expect to be using decedent samples based on more than one year's deaths, so as to explore what might be more appropriate models.

In any case, standard errors can be calculated according to the above model (using the pseudo-replicate balanced half-sample procedures of McCarthy [11]). While the formulation gives us six degrees of freedom, we could only afford to use one pair of "replicates" (the cost per pseudo-replicate was about $750); therefore, the standard errors had to be based on an estimator with only one degree of freedom.

4/ The social security taxable maximum was $3,600 in 1951; by 1971, however, the maximum had risen to $7,800.

REFERENCES

[1] Atkinson, A. "The Distribution of Wealth in Britain in the 1960's--the Estate Duty Method Reexamined," Studies in Income and Wealth, National Bureau of Economic Research, no.39, 1975, pp. 277-319.

[2] Langley, K. "Discussion: Wealth in Britain--the Estate Duty Method Reexamined," Studies in Income and Wealth, National Bureau of Economic Research, no. 39, 1975, pp. 321-328.

[3] Smith, J. and Franklin, S. "Expost Estimations of Weights in a Nonprobability Sample," paper presented at the May 5, 1976, Washington Statistical Society Methodology Session on Validating the Estate Multiplier Technique.

[4] Scheuren, F. "Historical Perspectives on IRS Wealth Estimates with Prospectives for Improvements," paper presented at the May 2, 1975, National Bureau of Economic Research Workshop on Estate Multiplier Estimates.

[5] Internal Revenue Service. Statistics of Income--1972, Personal Wealth Estimated from Estate Tax Returns, Washington, D.C., U.S. Government Printing Office, 1976.

[6] Natrella, V. "Wealth of Top Wealthholders--1972," 1975 American Statistical Association Proceedings, Business and Economic Statistics Section.

[7] Smith, J. and Calvert, S. "Estimating the Wealth of Top Wealth-holders from Estate Tax Returns," 1965 American Statistical Association Proceedings, Business and Economics Statistics Section.

[8] 1974 Social Security Bulletin Annual Statistical Supplement, pp. 69-77.

[9] Scheuren, F. and Oh, H.L. "Some Preliminary Synthetic SSA Cohort Survival Rates, 1950-1971," Unpublished SSA working paper, April 1976.

[10] Finch, R. "Examples of Sampling Variability in the Continuous Work History Sample," paper presented at the May 13, 1976, CWHS Users Conference.

[11] McCarthy, P. "Replication: an Approach to the Analysis of Data from Complex Surveys," Vital and Health Statistics, National Center for Health Statistics, series 2, nos. 14 (1966) and 31 (1969).

[12] Rosenwaike, I. Seasonal Variation of Deaths in the United States, 1951-1960, Journal of the American Statistical Association, vol. 61, pp. 706-719.