

THE INTERGENERATIONAL WEALTH STUDY: PROSPECTS FOR DATA ANALYSIS AND  
METHODOLOGICAL RESEARCH

Mary F. Bentz, Internal Revenue Service\*

This paper describes the plans for the creation and future use of a linked intergenerational United States Estate Tax return archive that is planned to encompass all the returns filed since the inception of the U.S. Estate Tax in 1916.

U.S. Estate Tax returns are filed for the estates of individuals who at death held gross assets exceeding an annual legal filing threshold. This threshold has been revised upward since 1916 in order to ensure that only the wealthiest segment (generally the top five percent or less) of the population is subject to filing a return (Figure 1). The following information is required to be reported on the return:

- assets
- liabilities
- estate tax paid
- identification of heirs
- amounts of bequests

In addition, data on the social and demographic characteristics of the decedent are provided, thus enabling a fairly accurate assessment of a person's socioeconomic status.

The estate archive, to be made up of economic, demographic, and social data from estate tax returns, will open up areas of research heretofore little or completely unexplored [1]. Using this data base, historical and contemporary analytical applications can be made involving such areas as the operation of the estate tax, the composition of asset holdings, and the characteristics of elites. The archival data can also be linked with information from other sources such as gift and income tax returns, probate records, death certificates, and social security records for research applications in a variety of disciplines.

Most interestingly, the archive will include identifying information on heirs and bequests that will permit the linkage of individuals in successive generations related by family and/or economic ties. This information will be used to trace the transfer of wealth from decedents to beneficiaries over several generations.

This project is especially intriguing in that it provides opportunities for both data analysis and methodological experimentation. In fact, while our completed archive will certainly be provocative, the actual process of creation, linkage and dissemination of the archive promises to be just as enlightening and probably even more challenging.

TAX RESEARCH

As part of an ongoing research program designed under the direction of the Office of Tax Analysis (OTA) in the Office of the Secretary of the Treasury, the Statistics of Income (SOI) Division regularly produces statistical studies of estate taxation [2]. These data, which have been published since 1922, provide a historical view of the estate tax and its effect on the accumulation and composition of wealth. The data are used by OTA to study the characteristics of the population subject to the tax, to assess the impact of existing and proposed laws on the tax revenues, and to model such relationships as the allocation of estates between charitable and noncharitable bequests.

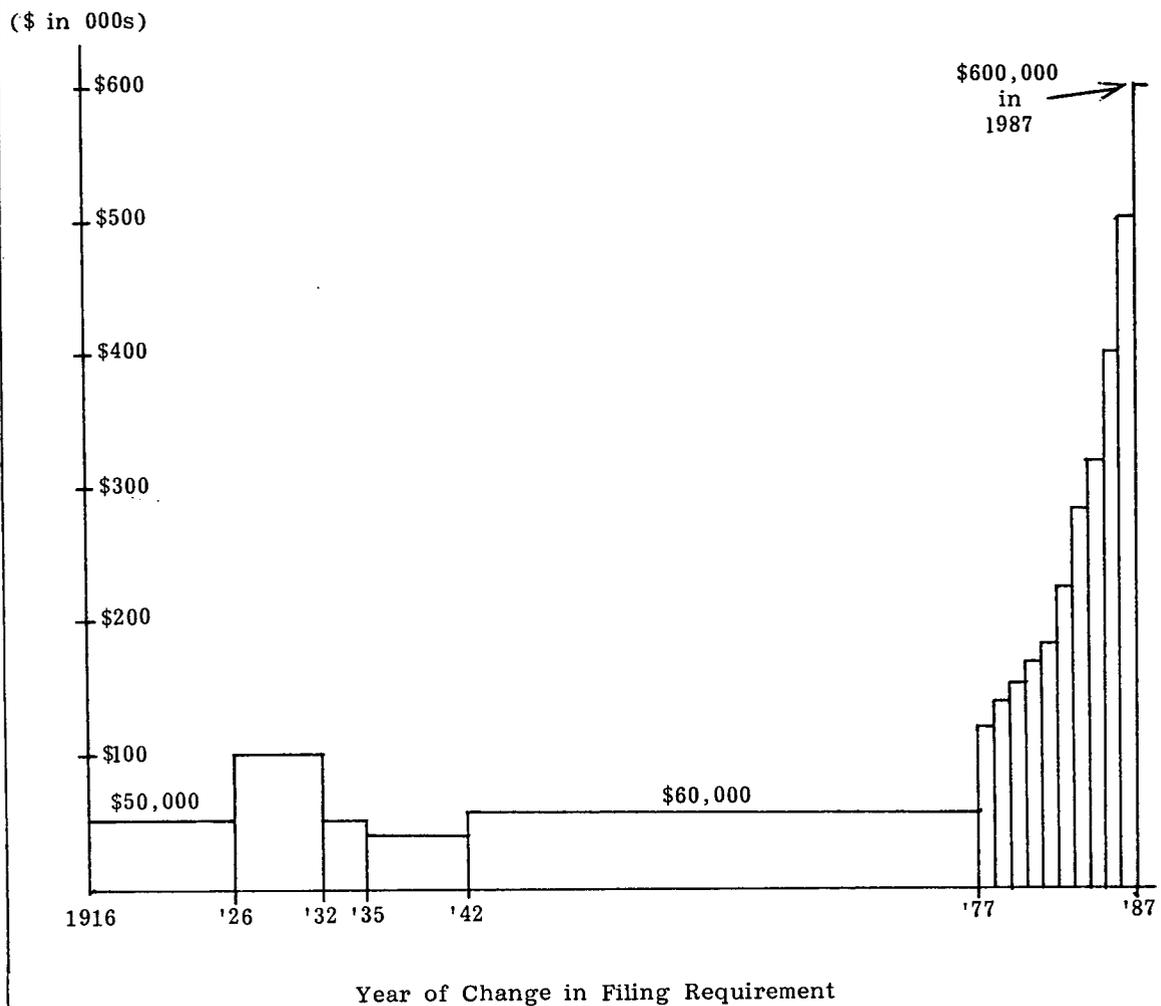
The Estate Tax file is also used as a basis for the SOI Estate-Income Tax Collation study and the Gift Tax study. The Estate-Income Tax Collation file provides an opportunity for research on the relationship between income and wealth by matching the income tax returns for the decedent and for each beneficiary of the estate with the estate tax return [3]. For decedents, returns for the year of death and the two preceding years, and for beneficiaries, the returns for the year prior to and three years after the decedent's death are associated with the estate tax returns. The sample permits comparisons of the wealth and income levels of decedents with the income levels of their beneficiaries, both before and after the year of death of the decedent or at the point at which the transfers of wealth occurred.

Our current Collation Study, which will be completed in 1986, will add information from fiduciary income tax returns and gift tax returns to the estate and individual income tax return data. The following areas will be given special attention: the timing and nature of gifts, the use of various trust instruments to reduce the estate tax burden, donor/donee relationships, and the interaction of lifetime income, saving, wealth, and wealth transfers. The file also provides a means for examining the relationship between income realized from capital and the value of the assets that produced that income.

WEALTH RESEARCH

A time series of the distribution of U.S. wealth based on year of death will also be produced from the estate tax return data. Using the Estate Multiplier technique, estimates of the wealth of the living population can be derived by applying multipliers based on

Figure 1.--Estate Tax Return Filing Requirements, 1916-1987



mortality rates to data reported on the returns. Decedents for whom returns are filed can be seen as a stratified sample of the living population for a given year: the sample is drawn by death and the sampling rates are simply the mortality rates associated with specific age/sex categories.

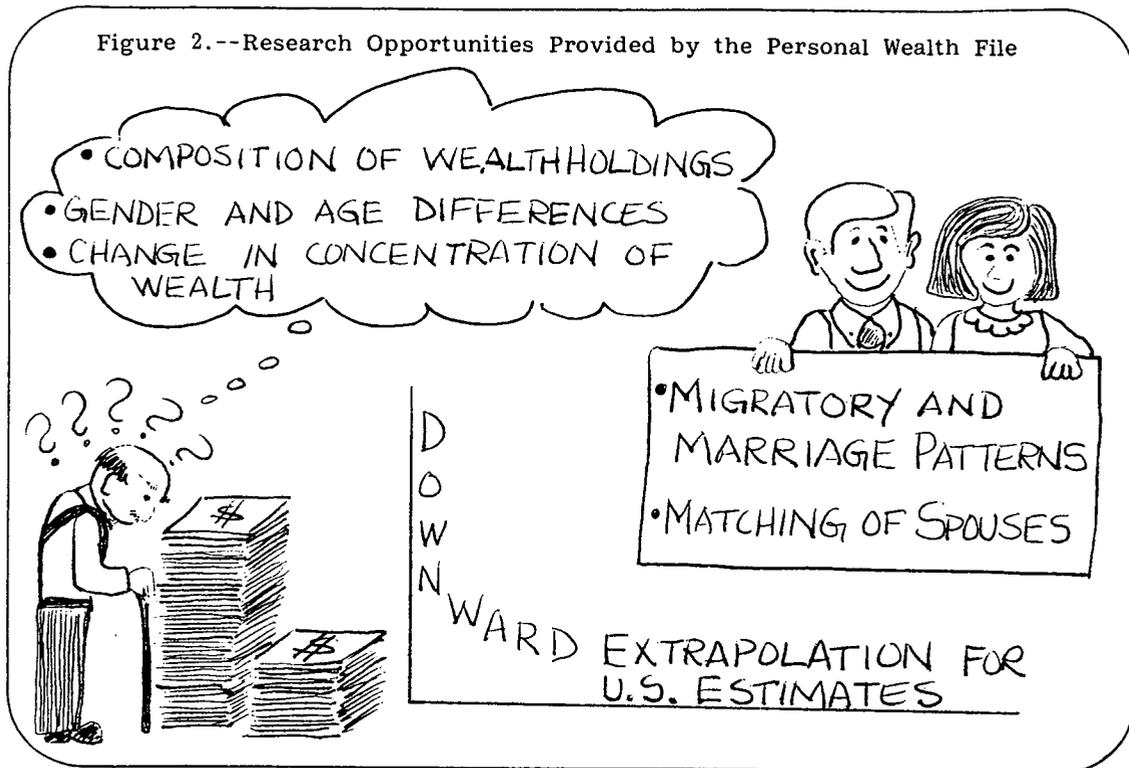
The personal wealth file will present considerable opportunities for research in the composition of wealthholdings and the share of total U.S. assets held by the wealthy; migratory and marriage patterns of the wealthy; and gender and age differences in asset holdings and changes in these differences over time. While our estimates of personal wealth have thus far been limited to individuals, household estimates can easily be produced by "marrying" individuals based on sex, age, and marital status included on the file. Spouses can also be matched within and across years to produce an even more accurate picture of a couple's wealth.

Moreover, estimates of the distribution of wealth of the population as a whole can be extrapolated downward from the estimates produced for the top wealthholders, especially in conjunction with outside information that could be added to the archival file (Figure 2).

In light of the difficulties encountered by wealth researchers over the years in producing a distribution of the wealth of the entire U.S. population, such a comprehensive model could prove to be extremely valuable. The major weakness in other research relates to estimates for top wealth-holders: surveys of the wealthy are subject to high nonresponse, inaccurate estimates, and the omission of assets, while capitalization of income techniques ignore both debts and assets that do not produce income.

Perhaps the most unusual contribution of the archive to wealth research will be the

Figure 2.--Research Opportunities Provided by the Personal Wealth File



inclusion of identifying information on heirs and bequests that will permit the linkage of individuals in successive generations related by family and/or economic ties. This information will be used to trace the transfer of wealth from decedents to beneficiaries over several generations (Figure 3). The linked intergenerational database will provide opportunities for studying the role of inheritance in wealth accumulation, and the pattern of wealth transfers. Of particular analytical interest will be the information gathered on the role of trusts in the intergenerational transfer of wealth.

OTHER RESEARCH

The proposed archive alone offers a great deal of information that could be useful to researchers in such disciplines as economics, demography, law, epidemiology, and sociology. In addition to the usual asset, debt, tax and bequest data, the following information have been required to be reported on various revisions of the estate tax form:

- occupation
- date and length of widowhood
- number of children
- cause of death
- length of last illness
- citizenship
- place of birth
- date of marriage to surviving spouse
- date of divorce.

For returns where these items have not been requested or reported, other sources of these

Figure 3.--Contribution of the Archive File to Wealth Research

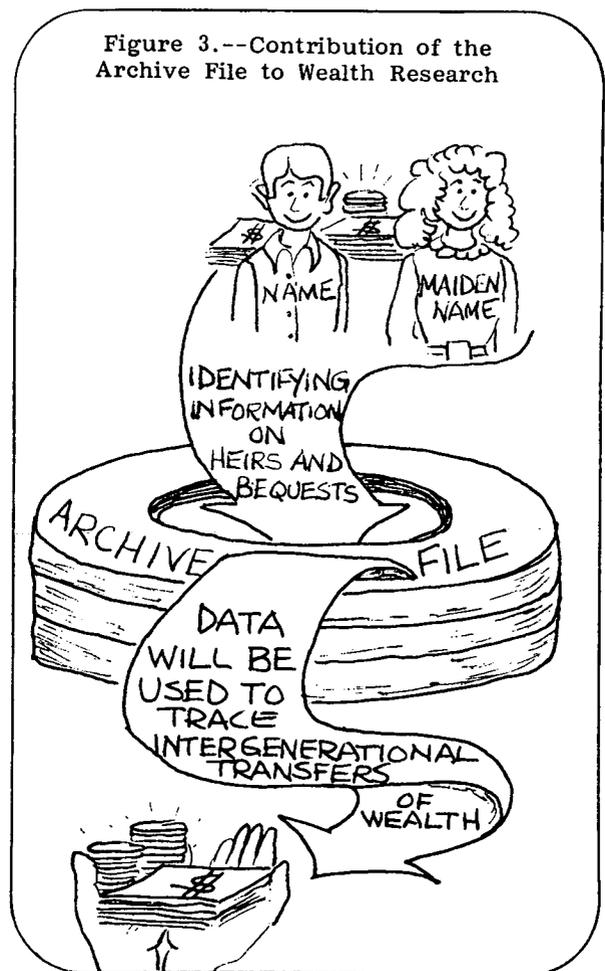
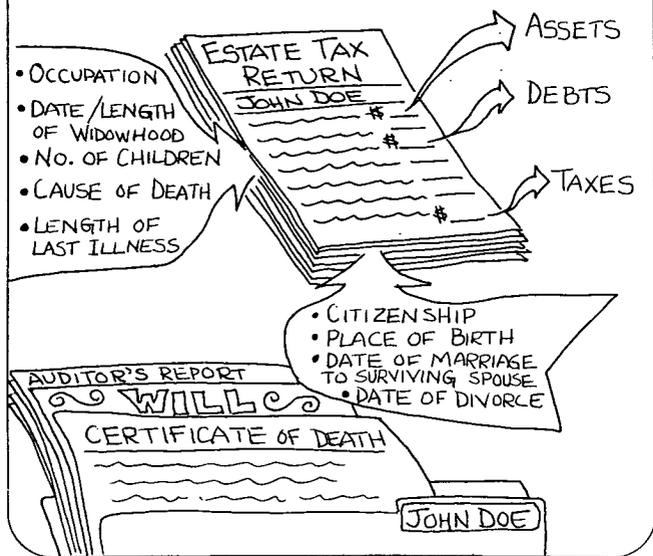


Figure 4.--Sources of Information for the Estate Archive File



data are available. The tax return itself is supported by a copy of the will, trust instruments, appraiser's reports, auditor's reports and, in many cases, the death certificate. Information on ethnic origin, length of last illness, cause of death, and nationality are all reported on the death certificate (Figure 4).

Additionally, once the estate archive is established, it will be possible to add data from other files to the estate records. Probate records could be used to assess the net worth of heirs who at death held assets less than the estate tax return filing threshold. Social security earnings records could be traced for different periods in the life of an individual, thus providing insights into the relative importance of lifetime earnings versus inheritance in wealth accumulation. Income and gift tax returns could be used, as in current SOI research, to examine the interaction of the estate, gift and income tax in a model of lifetime utility maximization. Death records for the decedents can also be added: SOI is currently pursuing a small-scale project that matches death certificates to the corresponding estate tax returns. Information on occupation and cause of death reported on the death records will then be used to study socioeconomic differentials in mortality and morbidity (Figure 5).

An added benefit of our planned mortality research is that the new wealth-specific mortality rates will also help to perfect the mortality differentials used in the Estate Multiplier technique.

#### METHODOLOGICAL ISSUES

The steps outlined below represent an

idealized construction of our long-range plans for the project:

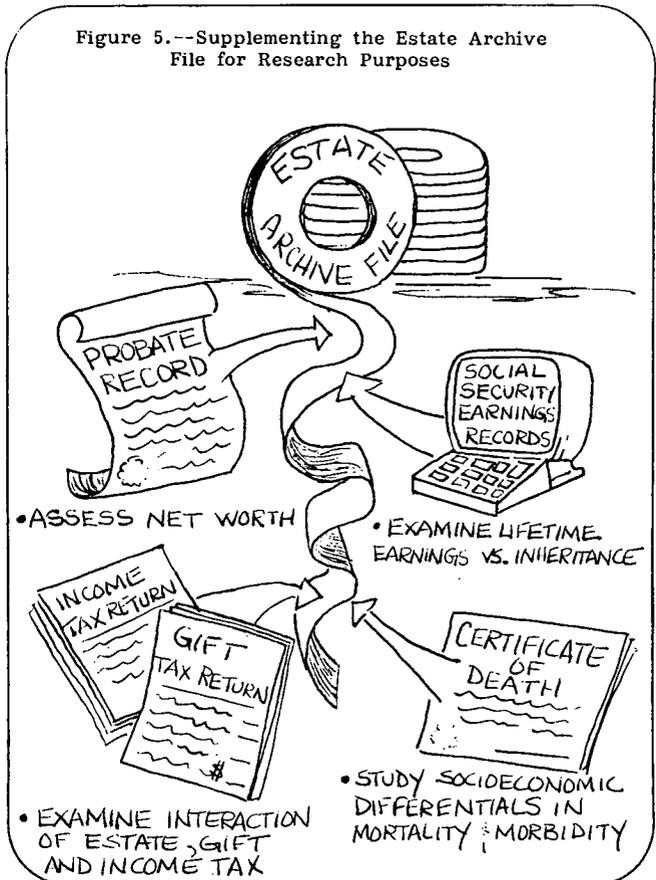
- creating an archive of economic, social, demographic, and legal information extracted from the approximately four million federal estate tax returns filed since 1916;
- developing an intergenerational database linking individuals related by family and/or economic ties; and
- developing a means of scholarly access to the archive.

#### Creation of an Archive of Estate Tax Data

All returns for all the years 1916-1945 (about 360,000) will undergo a limited data retrieval process in order to establish a machine-readable set of information that will permit the linking of returns of family members in successive generations [4]. Limited information enabling the computation of total assets and net worth for a time series of wealth estimates will also be extracted from each return.

In addition to the basic data to be extracted from all returns, detailed information on the composition of assets will be extracted from about 145,000 returns filed in selected

Figure 5.--Supplementing the Estate Archive File for Research Purposes



years. Returns for decedents who died in the following years will be subject to this complete (extra) data capture: 1916 to 1920, 1928 to 1930, 1938 to 1940, 1944, 1948 to 1950, 1953, 1958 to 1960, 1968 to 1970, and 1978 to 1980. Groups of three years have been selected to center around years ending with a 9 (e.g., 1928 to 1930) in order to coincide with U.S. Bureau of the Census data on income, which is available for years ending in a nine. Wealth estimates for the single years 1944, 1953 and 1958 were designated so that comparison can be made to wealth estimates produced by Horst Mendershausen for 1944 [5], by Robert Lampman for 1953 [6], and by Jim Smith for 1958 [7].

For each year not subject to a complete data capture, detailed asset information will be obtained for the estates of decedents with assets greater than the equivalent of \$5 million in 1982 (but no less than 500 decedents annually). This information will provide a picture of the concentration of the wealth of the most affluent individuals over the years. These returns will also be microfilmed so that they will be available for future use.

#### Developing a Linkage Methodology

As part of the intergenerational study of wealth transfers, linkage of beneficiary data from estate tax returns with the estate tax returns eventually filed for each decedent's beneficiaries will be ongoing as each year's data abstraction is completed. As each annual estate file is produced, a sub-file of all individuals listed as beneficiaries on estate tax returns will be created. Identifying information from this beneficiary file will then be matched to decedent files for the current year and following years to determine if the heir has died and left an estate of his or her own. The intent is to match as many heirs as possible with the estate tax returns subsequently filed for them using the identifying information from the beneficiary file and from the subsequent estate tax file. After the initial computer matches are made, a more intensive manual search, possibly using information obtained from social registers and obituaries, will be conducted for the remaining unmatched heirs, especially those expected to leave estates of their own. For more recent years, it may also be possible to use death records and social security files to verify the death of unmatched heirs.

The matching of decedents and heirs over time will be complicated by several factors. First, social security numbers were not reported on the returns for decedents until 1961, and for beneficiaries until 1965. For more recent years, the presence of social security numbers will facilitate the search substantially. However, much of the linkage of records of beneficiaries with their eventual estate tax returns will be based solely on the following information: name, geographic location, and age of the beneficiary at inheritance.

Secondly, name, which will necessarily be an important matching variable, is subject to considerable reporting and spelling variations. Name changes, especially for younger females, also introduce a degree of uncertainty. Some solutions to be explored are pre-processing of files to standardize names, blocking of surnames by phonetic codes, and multiple matching techniques.

Finally, measuring the error associated with the record linkage will be difficult because we do not know the number of beneficiaries who will hold enough assets at death to make their estates subject to filing the estate tax return.

The project will provide an opportunity to test different software packages for use in matching, and a chance to evaluate and compare existing record linkage systems.

#### Developing a Means of Scholarly Access to the Archive

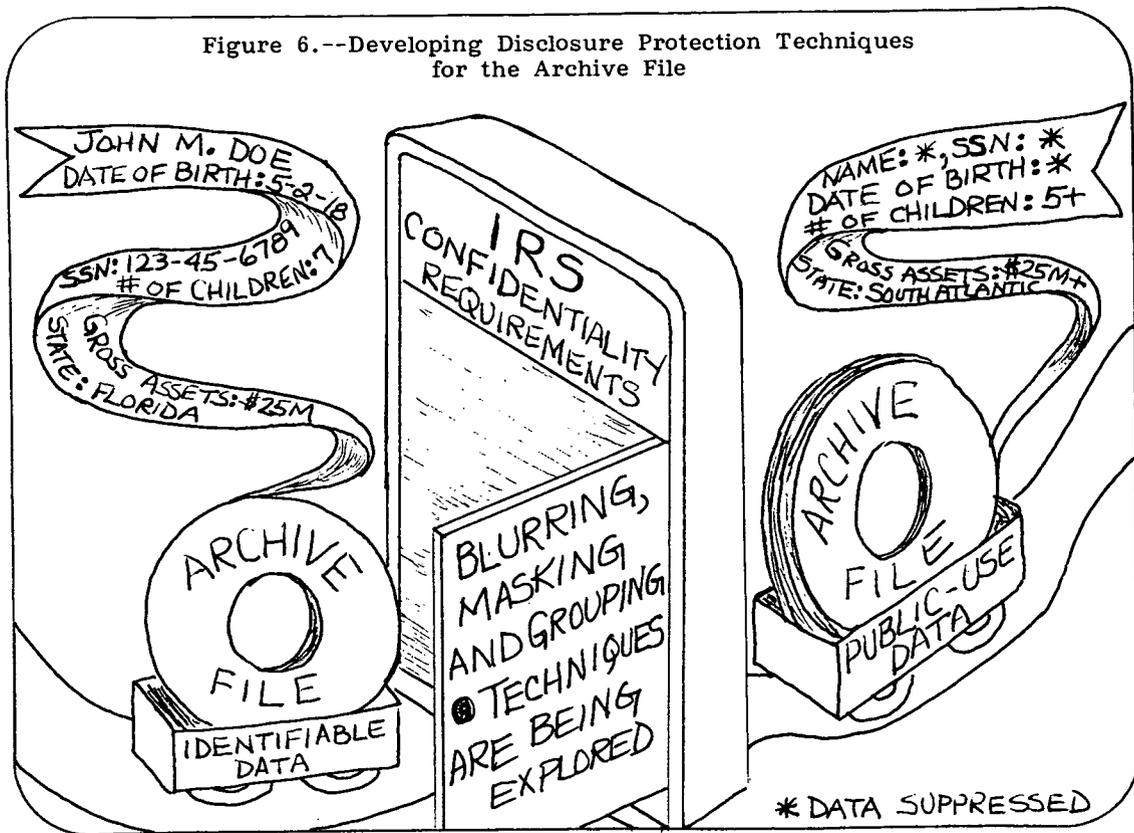
In order to make the archival file publicly accessible, it will be necessary to develop protections against the disclosure of the identity of individuals on the file. Information that could be linked with high probability to a particular individual must also be suppressed. The existence of externally available information that could be linked to a data set makes disclosure protection even more difficult. We are currently exploring various blurring, masking, and grouping techniques that would satisfy IRS confidentiality requirements while preserving the usefulness for outside researchers of the archival public-use data file (Figure 6). Other possibilities are a data-use facility at IRS, establishment of a parallel facility outside IRS, and joint statistical agreements with individual scientists. Our strategy will be to eliminate as little information as is necessary to ensure the confidentiality of the data.

#### PROJECT STATUS

As of June, 1986, approximately 187,000 estate tax returns filed for decedents who died from 1916 through 1935 were processed at the IRS Service Center in Ogden, Utah. During this phase, the data obtained from the returns included limited identifying and demographic information as well as information on the size of gross estate, amount of tax, and other information enabling estimates of total assets and net worth. Detailed asset information was extracted from returns filed in selected years and from returns with a gross estate valued at the equivalent of \$5 million in 1982. Using these returns, we have created a file from which we are currently in the process of producing some general tabulations.

At present, we plan to produce three public-use data files. One will be a matched file of decedents and beneficiaries. The other will be a complete archival file, broken into

Figure 6.--Developing Disclosure Protection Techniques for the Archive File



ten-year blocks. From the archival file, we will also produce an "asset" file, including just those returns for which complete asset information is recorded. These three files will be protected from disclosure by one or more of the techniques mentioned above. Additional public-use files will be masked and provided to researchers on a case-by-case basis to the extent possible, given IRS confidentiality policies. Alternatively, it may be possible to create public-use files that can be easily modified according to users' needs.

NOTES AND REFERENCES

\*Mary Bentz is no longer with the Internal Revenue Service. For further information on this study, contact Marvin Schwartz, Statistics of Income Division.

- [1] Smith, James D., Socio-Economic Structure and Dynamics of American Households in the 20th Century, The Institute for Social Research, University of Michigan, 1984.
- [2] As mandated in the Internal Revenue Code, the IRS has regularly conducted statistical studies of estate and gift taxation since the inception of these taxes. Also, beginning with 1962, estimates of the personal wealth of top wealthholders have been produced as a by-product of the estate tax return file, using the Estate Multiplier technique. SOI now provides estate tax

return data on a filing year basis and quadrennial estimates of wealth for specific years.

Reports of estate tax return filings and preliminary personal wealth estimates based on the returns filed each year are published annually in the Division's Statistics of Income Bulletin. For the most recent articles, see Schwartz, Marvin, "Estimates of Personal Wealth, 1982," Statistics of Income Bulletin, Winter 1984-85; and Bentz, Mary F., "Estate Tax Returns, 1983," Statistics of Income Bulletin, Fall 1984.

In addition, statistical reports documenting various ongoing research projects are produced periodically for the annual American Statistical Association (ASA) meetings. See Bentz, Mary F. and Schwartz, Marvin, "Continuing IRS Estimates of Personal Wealth in the United States," Statistics of Income and Related Administrative Record Research, Internal Revenue Service, 1984. Also published in 1984 American Statistical Association Proceedings, Survey Research Section, pp. 384-389.

A compendium of wealth research studies, including also the results from the Estate/Income Tax Collation study, will be published in 1987.

- [3] Steuerle, Eugene, "Wealth, Realized Income, and the Measure of Well-Being," Horizontal Equity, Uncertainty, and Economic Well-Being, David, Martin, and Smeeding, Timothy (editors), National Bureau of Economic Research, University of Chicago Press, 1985, pp. 91-124.
- [4] For years after 1945, we plan to process only returns reporting total assets in excess of an amount equivalent to \$300,000 in 1982 dollars. (Beginning in 1982, the ongoing Estate Program has been designed to be compatible with this archival effort.)
- [5] Mendershausen, Horst, "The Pattern of Estate Tax Wealth," Part III of Raymond W. Goldsmith's A Study of Saving in the United States, Princeton: Princeton University Press, 1956, pp. 277-381.
- [6] Lampman, Robert J., The Share of Top Wealthholders in National Wealth, 1922-56, Princeton: Princeton University Press, 1962.
- [7] Smith, James D., The Income and Wealth of Top Wealthholders in the United States, 1958, Dissertation, University of Oklahoma, 1966.
- [8] For the most recent Bulletin articles see Schwartz, Marvin, "Estimates of Personal Wealth, 1982," Statistics of Income Bulletin, Winter 1984-85; and Bentz, Mary F., "Estate Tax Returns, 1983," Statistics of Income Bulletin, Fall 1984.
- [9] Bentz, Mary F. and Schwartz, Marvin, "Continuing the IRS Estimates of Personal Wealth in the United States," Statistics of Income and Related Administrative Record Research, Internal Revenue Service, 1984. Also published in 1984 American Statistical Association Proceedings, Survey Research Section, pp. 384-389.