THE 1998 GIFT TAX PANEL STUDY: USING THE IRS RETURNS TRANSACTION FILE AS A SAMPLE FRAME

Martha Britton Eller and Tamara L. Rib, Internal Revenue Service
Martha Britton Eller, P.O. Box 2608, Washington, D.C. 20002

Key Words: Stratified random sample, Data cleaning, Data uses and limitations

Introduction

The Federal gift tax is one of three taxes included in the U.S. transfer tax system, which, simply stated, is a unified system that taxes transfers of property completed both during life and at death. The two other components of the U.S. transfer tax system are the estate tax, applied to the value of property transferred at death, and the generation-skipping transfer tax, applied to the value of property transferred to trust for the benefit of an individual or individuals two or more generations below that of the grantor, or donor.

The first Federal gift tax was introduced in the Revenue Act of 1924. Congress imposed the 1924 tax after it realized that wealthy Americans could avoid the estate tax, introduced in 1916, by transferring wealth during their lifetimes, called *inter vivos* giving. Tax-free *inter vivos* gifts effectively negated the estate tax's capacity to redistribute wealth accumulated by large estates and removed a source of revenue from the Federal government's reach (Johnson and Eller, 1998).

The first gift tax was short-lived. Due to strong opposition against estate and gift taxes during the 1920s, Congress repealed the gift tax with the Revenue Act of 1926 (Zaritsky and Ripy, 1984). Reintroduced in the Revenue Act of 1932, when the need to finance Federal spending during the Great Depression outweighed opposition to gift taxation, the 1932 gift tax allowed a grantor to transfer $50,000 during his or her life and allowed a $5,000 annual exclusion per gift recipient, or donee. The 1932 Act set gift tax rates at three-quarters of the estate tax rates, a level maintained until 1976, when Congress passed the Tax Reform Act (TRA) of 1976 and created the unified estate and gift tax framework that consisted of a "single, graduated rate of tax imposed on both lifetime gift and testamentary dispositions" (Zaritsky and Ripy, 1984). The generation-skipping transfer tax was also introduced in TRA of 1976.

During the years since 1932, features such as the marital deduction and rules on split gifts were introduced to gift tax law, but the predominant changes to the law were adjustments to the amount of lifetime exemption and annual exclusion. A gift is taxed under the law that is in effect during the year in which the gift is completed, or given. According to transfer tax law in effect for gifts completed in 1997, the focus of this paper, a grantor was required to file a Federal gift tax return (Form 709) for transfers of property in excess of $10,000 per donee, and the lifetime unified credit—equal to the tax on the lifetime-giving threshold for 1997, $600,000—was $192,800. Under Internal Revenue Code (IRC) section 2511(a), the gift tax applies to a broad spectrum of gifts, "whether the gift is in trust or otherwise, whether the gift is direct or indirect, and whether the property is real or personal, tangible or intangible." Regulation 25.2511-1(c)(1) provides that a completed gift, one that is subject to tax, is "any transaction in which an interest in property is gratuitously passed or conferred upon another, regardless of the means or device employed."

Gift tax data extracted from Federal gift tax returns provide a glimpse into the economic behavior of predominantly wealthy Americans. Such behavior includes donors' transfer of money and other assets to gift recipients and the creation and continued funding of trusts, both of which are reported on gift tax returns. Since individuals are required to file annual returns for gifts completed during a prior calendar year, it is possible to construct a panel of gift tax returns filed during life for a subset of U.S. taxpayers, thereby capturing the lifetime giving patterns exhibited by the group.

The Statistics of Income Division (SOI) of the Internal Revenue Service (IRS), an organization that extracts and publishes data from Federal tax and information returns, initiated the 1998 Gift Tax Panel Study in order to examine gift tax revenue, as well as the lifetime giving patterns of wealthy Americans. At the close of the study, SOI will have obtained and extracted data from post-1976 returns filed by donors included in the study, creating a retrospective panel of returns for selected donors. Resultant data will facilitate the research of lifetime giving patterns and patterns of trust creation and maintenance, among other goals.

The 1998 Gift Tax Panel Study is an exception to the usual design of SOI studies in which statistical samples are based on estimates of given populations of returns. Because SOI sampling of returns normally occurs immediately after IRS processing of returns for tax revenue purposes, the final population of returns is not known at the time of

sample design and weekly selections. But, the population of gift tax filers was known before the inception of the study, because the sample frame for the study was the 198 IRS Returns Transaction File (RTF), a data file that contains all tax year 1997 gift tax returns that posted to the IRS Master File during revenue processing in 1998.

This paper will discuss the RTF and its use as a sample frame in SOI's statistical study of gift taxation. It will address issues of data cleaning, sample design, weighting, imputation and data uses and limitations.

The 1998 Returns Transaction File (RTF)

The IRS Returns Transaction File (RTF) is a data file that contains records for returns processed during a calendar year by the revenue processing function of IRS. It is a subset of the data in the IRS Master File. With few exceptions, information entered on the returns processed by IRS, regardless of return type, is available, in abbreviated form, on the IRS Master File and RTF. The 1998 RTF for Federal gift tax returns included records for all tax year 1997 and earlier gift tax returns processed by IRS during filing year 1998, regardless of the year in which the gift was given. Since applicable gift tax law is determined by the year in which a gift is given, and since the majority of gifts given in one year are reported in the following year, a single gift year, 1997, was chosen as the focus year for the 1998 Gift Tax Panel Study.

Prior to sample selection, SOI analysts excluded amended returns, duplicate returns, out-of-scope returns, and records that merely represent transactions, not returns (i.e., "invalid" records), from the RTF. Amended returns adjust returns previously filed and, in many cases, are simply supplements to original returns. As such, amended returns usually are incomplete. About 0.6 percent of returns included on the original file of 239,985 returns were amended and, therefore, removed. The file also included records for duplicate returns filed for gift year 1997. Duplicate returns were reviewed and ordered by date of IRS receipt, and only the first return, the one with the earliest date, was retained. About 1.8 percent of returns on the original file were duplicates. Returns with zero and negative values for a variable of interest—size of total gifts—were considered out-of-scope and excluded from the file. About 5.1 percent of returns on the original file met this criterion. Any records that the IRS defined as "invalid" were also excluded. Invalid records typically correct a transaction on a previous record and do not themselves represent a return. The IRS assigns a zero prefix to the social security number (SSN) on invalid records. About 1.1 percent of records on the original file were invalid. After cleaning the RTF, the final population of gift tax filers for filing year 1998 (1997 gifts) was 219,414. These returns became the sample frame for the study.

While the RTF for gift tax returns contains the population of filer records and includes many of the variables used in the computation of tax and in the calculation of total taxable gifts, there are problems in relying solely on RTF data for population estimates of these variables. One persistent problem that SOI analysts encounter when working with the RTF is that some arithmetic relationships between variables for a given record are not correct. And, as found in the course of the study, the stratifiers, taxability status and total gifts, were incorrect in several instances on RTF taxpayer records.

In addition to the uncertainty in the accuracy of the RTF data, another problem is that important pieces of information are not available on the file. Such information is only available on the Federal gift tax return itself: the size and type of gift, as well as the name of the gift recipient, whether an individual or a trust. If the gift recipient is an individual, there is evidence for deducing the sex of the individual and the individual's relationship to the donor. Similarly, if the gift recipient is a trust, the type of trust, whether marital, family, insurance, etc., is also available. These donee and gift data are important to SOI's data customers, and, without SOI personnel extracting such data from gift tax returns, they would not be available to customers. Overall, then, SOI-edited data provide more accurate and detailed information on donors, donees and gifts. As noted earlier, the panel feature of the study provides further information on areas of interest to customers, such as patterns of giving, and trust creation and funding.

Sample Design for the 1998 Gift Tax Panel Study

The sample frame for the 1998 Gift Tax Panel Study included 219,414 Federal gift tax returns filed for gifts completed in 1997. Based on budget and other constraints, a target sample of 10,000 returns, or donors, was planned. SOI analysts met with data customers from the Office of Tax Analysis, the Joint Committee on Taxation and the IRS Estate and Gift Tax Administration in order to discuss possible data uses and to elicit ideas for the sample design. As a result of customers' input and SOI's analysis of the RTF, the final sample for the study was a random sample stratified by two variables: taxability status and size of total gifts (prior to the subtraction of annual exclusions and deductions in the calculation of total taxable gifts). Taxability

status is divided into two categories: nontaxable (i.e., no gift tax liability reported) and taxable (i.e., gift tax liability reported). The second stratifier, size of total gifts, is divided into four or five categories, depending on taxability status. Each stratum is labeled with a sample code.

Neyman allocation is used to assign the designated sample to the stratum. A Bernoulli sample is selected independently from each stratum. For nontaxable returns, sample rates vary from 0.9 percent, for returns with total gifts under $100,000, to 100 percent, for returns with $1 million or more in total gifts. For taxable returns, sample rates vary from 12.6 percent, for returns with total gifts under $100,000, to 100 percent, for returns with totals gifts of $1 million or more.

The sampling method for each stratum is based on the Taxpayer Identification Number (TIN), which is the donor's SSN, as found on the return and the RTF. First, a unique random number, called the Transformed Taxpayer Identification Number (TTIN), was calculated for each TIN. Then, the last four digits of the TTIN, a pseudo-random number, was compared to a range of numbers, based on a return's selection probability. If the number was less than the sampling rate multiplied by 10,000, the return was selected into the sample and processed. Any returns with total gifts of $1 million or more were automatically selected. The final sample included 9,914 Federal gift tax returns filed in 1998 for 1997 gifts. Because all post-1976 gift tax returns for each donor in the sample are included in the study, SOI estimates that the final panel will reach 50,000 Federal gift tax returns at the study's close.

Missing Returns

Because most Federal tax return documents are stored at IRS submission processing centers and Federal Records Centers across the country, it is almost inevitable that some of the documents in a sample are never found. Additionally, Federal gift tax returns are stored in individual taxpayer folders at the IRS submission processing centers. According to ideal storage procedures, all gift tax returns for a given taxpayer are stored together in a single donor folder. In most cases, folders are sorted and stored alphabetically by the taxpayer's last name. Of course, storage procedures vary among centers. For instance, some centers store gift tax returns in alphabetical order within an IRS district, an important organizational unit of the IRS. Other centers initially store gift tax returns by document locator number (DLN), the primary method of storing all tax returns filed at the centers during the year, then later in alphabetical order.

In reality, donor folders often do not contain all gift tax returns filed by taxpayers. In some cases, gift tax returns are simply placed in the wrong folder. In other cases, multiple folders for the same taxpayer exist at different centers for reasons that include taxpayer name changes (due to changes in marital status), changes in residency and IRS oversights. In addition, limited storage space forces centers to rotate documents, increasing the likelihood of misplaced or lost returns. Using the IRS Integrated Data Retrieval System (IDRS) to identify complete filing histories on taxpayers, it is possible, in theory, to locate all gift tax returns filed by donors selected into the sample, even if multiple folders across centers are created. However, if IDRS is not updated properly or timely, it may provide little help. For gift tax returns that are simply misfiled due to IRS handling errors, IDRS provides no help.

Personnel in the submission processing centers have utilized IDRS in the search for returns, and centers have worked together to consolidate all returns for each donor selected into the sample. In addition, SOI has worked closely with the IRS examination function in locating returns, since some gift tax returns included in the study may also have been selected for audit.

There are several ways to handle the missing returns, or non-response items, in the sample. Given that the sample was selected from a known population, most of a donor's information is known. There are current plans to impute the missing data using one or more imputation techniques on the previous or following year's data.

Base Weights and Imputation Methods

Each return in the sample will be weighted to reflect its share of the entire population. The base weight is computed by dividing the population count of filed returns in a given stratum by the number of sample returns in that same stratum. The weights are used to produce aggregate estimates for items of interest, such as total gifts, total deductions and total taxes.

In the event that the missing returns have not been located by the close of the study, missing data will be filled in with data available from the RTF. For missing 1997 gift tax returns, a record will be created using actual values from the 1998 RTF. This will provide available donor information. Gift recipient information will be copied from the closest prior year's gift tax return because these data are not provided on the RTF. For missing panel returns that were filed between 1988 and 1998, RTF data, available from 1988 to the present, will be used to duplicate the original return or fill in missing data

items. For missing returns filed prior to 1988, the average of values from the closest available year before and after will be substituted.

Future Plans and Conclusion

In January 2003, SOI will initiate a study of gift tax returns filed in 2003 for gifts completed in 2002. The new study will include a cross section of returns filed in 2003, as well as a sub-sample of returns selected in the 1998 study. This design will allow us to follow a panel of 1998 gift donors into the future. For the small sub-sample of 1998 donors, we will be able to extract data from returns filed between 1998 and 2003.

SOI analysts who worked on the 1998 Gift Tax Panel Study have learned much about the initiation and completion of a statistical study that uses the IRS Returns Transaction File as a sample frame. Use of the file affects almost every phase of the study, from sample design and selection to weighting. Some effects of using the 1998 RTF were positive, making the study easier to initiate and complete, while other effects were negative, creating obstacles to the study's completion. These effects revealed themselves as the study progressed. First, following extensive cleaning, which included the removal of amended, invalid, duplicate and out-of-scope records, the RTF provided 100 percent coverage of the gift tax filing population for 1998, an obvious positive effect of using the RTF as a sample frame. Second, access to population data for gift tax filers facilitated the research and sample design phases of the study, yet another positive effect. Because the RTF contains a population of historic filings, no matter how recent, its use as a sample frame requires the retrieval of returns after they have been filed, stored and, in some cases, audited. This third factor, the only negative effect, combined with the type of return and the way in which IRS controls it, introduces a greater possibility of missing returns, when compared to studies that sample returns as they are processed for revenue purposes. A positive, final effect, RTF data are available for filing years 1988 to present, so it is possible to use actual RTF values in place of missing values for those panel years.

IRS has also learned several lessons in the course of the 1998 Gift Tax Panel Study. U.S. taxpayers currently file Federal estate and gift tax returns in all 10 IRS submission processing centers across the United States. Beginning in January 2002, the IRS plans to consolidate the filing of estate and gift tax returns at the Cincinnati Submission Processing Center. This study's description of problems with storage procedures for gift tax returns may help IRS in its consolidation efforts. IRS files units have already begun to learn from their experience with the 1998 Gift Tax Panel Study and the search for historic returns. At least one submission processing center has entered all gift tax filings in an Access database for easier retrieval. This kind of inventory system, if introduced on a national level, may become invaluable to IRS, especially the IRS examination function, as it enters a new era in estate and gift taxation.

References

Johnson, Barry W. and Martha Britton Eller, "Federal Taxation of Inheritance and Wealth Transfers," *Inheritance and Wealth in America*, Robert K. Miller, Jr. and Stephen J. McNamee, eds. Plenum Press: New York and London: 1998.

Zaritsky, Howard and Thomas Ripy, *Federal Estate, Gift, and Generation Skipping Taxes: A Legislative History and Description of Current Law*, Congressional Research Service, Washington, D.C., Report No. 84-156A, 1984.

SOURCE: Statistics of Income Division, Internal Revenue Service, as presented at the 2001 Joint Statistical Meeting of the American Statistical Association, Atlanta, Ga. 2001.