Chapter 8: Improving the Precision of Sample Estimates

Origins of the Estate and Personal Wealth Sample Design

Paul B. McMahon, Jr., Internal Revenue Service

In Estates and Personal Wealth, we have two studies with different populations under consideration. The Estates Study is concerned with the assets, debts, and taxes left by a decedent who had more than a certain amount of wealth. The Personal Wealth Study, on the other hand, is focused on the wealth holdings of the living. For Estates, essentially all the population appears on a sampling frame, but, to study the living, we must rely on proxies that can be observed for only a portion of the distribution, the portion in the tail.

One set of samples is the source for the data in both series of studies.

We will first briefly describe the interest in these populations. The "questionnaire" in this set of surveys is an administrative record, the Form 706, *Estate Tax Return*, and the sampling frame is a system of electronic records derived from the initial filing. We will provide a bit of background on these as well.

We focus on the studies initiated since 1982, with strata designs that changed somewhat over that time. While some previous papers have addressed certain estimation issues, such as with the Personal Wealth Estimation (Johnson and Woodburn, 1994), there have been only the briefest descriptions of the strata design or concepts.

Our goal, then, is to show how the different requirements for studies of the two populations affect this one sample design, and how that design has evolved in the light of tax law changes.

Finally, we will discuss some future directions for the series, in light of pending legislation.

▶ Analysts and Uses

The two main sponsors of these studies are the Office of Tax Analysis in the Department of the Treasury and Congress's Joint Committee on Taxation. Their objective is to gather data for oversight on the operation of the tax laws and, in this case, on Estate Taxes, and projecting the effects of proposed changes to those laws. This is not limited to the revenue aspects of the tax laws.

That is, this study has to meet two uses. First, the measurement of current law, and second, determining the effect on the living population who have estates large enough for the eventual filings. In order to look at trends in the analysis, we need to be concerned about the effect of economic conditions at the time of the observations (the date of death), the time of life considerations (youthful spenders versus middle-age savers, for example), and what the sociologists call age cohorts, where history affects economic decisions (the Depression generation's thrift).

There is also an underlying philosophical question: Does the operation of the Estate Tax, in concert with a graduated income tax, prevent the concentration of wealth into few hands? At the beginning of the twentieth century, some politicians, like Theodore Roosevelt argued in favor of the Estate Tax on just this issue. More recently, there have been numerous articles this past spring in the *New York Times* and the *Wall Street Journal*, for example, on the concentrations of incomes. Income is often taken as a proxy for wealth; so, this question is clearly of continued interest.

Indeed, using data from Estate Tax Returns dating back to 1916, the National Bureau for Economic Research (NBER) published a working paper that considers this very concentration issue (Kupczuk and Saez, 2004). Although the data used in that study are from many years in the past, the sample designs for most of those years actually originated in the mid-1980's and reflect the plans developed for sampling more recent tax filings.

▶ The Administrative Records

The basic data for these studies use the records that arise from what some have called the "Death Tax." It is more accurate, though, to call it a transfer tax, as the

change of an asset's title to some beneficiary or heir is the proximate cause for this tax or its complement, the gift tax. The tax return, which acts as the questionnaire for our studies, is Form 706, *Estate Tax Return*.

The assets that are considered for this tax are everything owned by the decedent: art, bonds, cars, personal effects, through to zoom lenses and beyond. That is, the filing is based on a complete inventory of an individual's possessions. In this, it is similar to the information that the Federal Reserve attempts to obtain in its Survey of Consumer Finance.

There are major differences between the data collected for the Federal Reserve surveys and the IRS studies, however. First, the tax form also includes insurance payments to the estate and gifts made before the decedent's death, which would not be included in the Finance Survey. Then, the law permits deductions for the costs of such items as estate administration, the funeral, and legal counsel, as well as exempting the contributions to charities and the spouse of the decedent.

Another difference is that the value of the assets is usually assessed at the time of death, not as of some common reference date for all respondents.

The main difference, though, arises from the populations these two sets of studies targets. The Survey of Consumer Finance seeks to estimate the holdings of all households, while the Estates and Personal Wealth studies are limited to individuals who exceed a certain threshold set by the tax code.

Figure 1.--Estate Tax Return Filing Thresholds for Selected Years

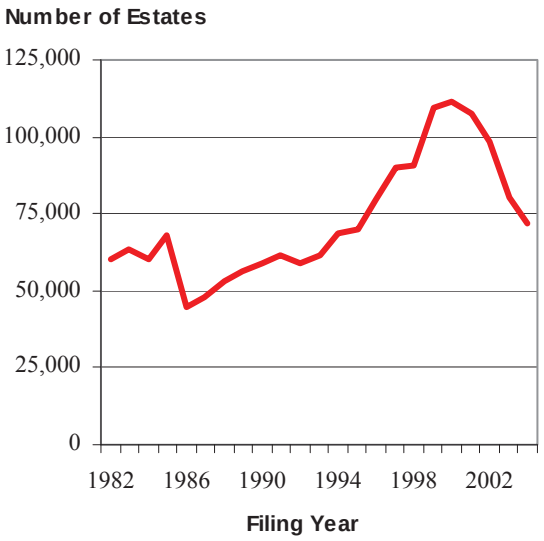
Year of Death	Gross Estate Threshold
1997	$600,000
1998	$625,000
1999	$650,000
2000 & 2001	$675,000
2002 & 2003	$1,000,000
2004 & 2005	$1,500,000
2006 – 2008	$2,000,000

If the value of those possessions at the time of the decedent's death is below the threshold amount shown in Figure 1, then there is no estate tax. That threshold varies depending on the year of the decedent's death. It is currently $1.5 million, rising to $2 million in January 2006. These values have been updated in the tax code periodically; in 1977, for example, the threshold was $60,000.

Filing is not required for smaller estates, though some do if the value is near the boundary. This may be due to the difficulty in itemizing all of an estate's assets. In those cases, amended returns will be filed, and perhaps a tax assessed, but such cases are outside the scope of this set of studies; we are only concerned with initial filings.

One can see the effect of raising the threshold quite clearly in Figure 2. In 1986, the exclusion was doubled, to $120,000, with a resultant sharp drop in filings and again, after the 2001 tax bill passed, which raised the limit several times in succession.

Figure 2.--Annual Filings of Estate Tax Returns



While the law and regulations provide one source of limitations on the studies, and thereby the design, another is in the physical properties of the documents and the processing regimen.

The Estate Tax Return is filed on paper as a large package with sections that are partly structured and partly

respondent-created. While Form 706 is, on the surface, highly standardized, the space allowed for some schedules (such as a list of heirs) is sometimes insufficient. This leads the attorney or executor to create substitute schedules of their own design.

The filing regulations also mandate the inclusion of the will, unless the decedent died intestate, appraisals of real property, and the death certificate. While the last may be relatively standardized, the will and appraisal(s) are not.

Moreover, all of these filings are subjected to an audit review, unlike the small proportion of Individual Tax Returns. Such audits keep the return unavailable for considerable lengths of time. Thus, the Statistics of Income studies must capture the return first and cannot wait for the entire population to become available; the sample must be selected as the returns are processed through the administrative pipeline.

The filing deadline for these documents is 9 months after the decedent's death. Extensions to this deadline are often required, because it takes time to locate some financial records, and for some assets to come to light. Since evaluating the effect of changes to the law is an objective, focus on a particular year of death means we must continue the selection over more than 2 years: the focus year and at least the following 15 months.

In practice, given the administrative environment, the minimum effective sampling period is 3 years. The additional months arise from the cycle of updating the computer programs, where the latest versions are introduced each January.

We want to use an electronic record in the sampling of these estates because, while selecting the returns as paper records ensures their retention for statistical purposes, this direct approach is costly and difficult and limits stratification options. The 1977 Study's manually-selected sample was limited to three strata, for example, and required considerable daily coordination with the ten national Service Centers where the returns were filed.

Yet the use of the computer records also gives rise to limitations. Ignoring audit trail codes, tracking data, and name and address information, there were only 16 amounts available in 1982, less than we can use today, but not by much. Most of those, 13, were involved in the calculation of the tax liability. This left a bare handful as possibly useful for sampling purposes, including some of the "code" fields.

Decedent's Year of Death was available. This was, and is, a tax-related field due to changes in the filing threshold; so, it was an administrative requirement.

For 1982, though, the Statistics of Income Division managed to convince the other interested parties within the Service that the age of the decedent could be useful. Rather than have a clerk calculate the age, though, the Service decided to include the Date of Birth. Gender, which could have been an important stratifier, is not available.

▶ The Stratifiers

Longitudinal studies in the sociology field have long noted that there are three effects to the group under observation: current events, time of life, and age cohort. We cannot easily address this last effect, that of the age cohort, at least not in the near future, because the observations on this group trickle in over such a long time.

We could address the aspect of current events' effect by focusing on all the decedents in a single year. "Current events," in this context, means not only the operation of economic conditions, but also the tax provisions then in force. Years ending in 2, 6, or 9 were selected; so, the first focus year included in this review is 1982.

Likewise, we could address the "time of life" through the age of the decedent (since we have the dates for both birth and death). This sociological concern has an economic component in the nature of financial holdings. For example, middle-aged people are often counseled to focus their investment strategy on growth, while retirees frequently look to revenue- producing equities. One tax consideration that arises is the unrealized capital gains included in the estate. By considering the age of the decedent, then, we can improve the measures in the composition of estates.

Age can also improve the reliability of the personal wealth estimates, which depend on this factor in the construction of the weighting classes.

Age and a focus year, though, would not aid in reducing the sampling error of the monetary estimates all that much, though. For that, we needed a variable that was reasonably correlated with the key amounts of interest. Given that this is a general sample to support ambiguous analysis (at the time of the design, anyway), that left Total Gross Estate as the monetary stratifier.

▶ Selection Method

Since the selection process was computerized, we took advantage of a Bernoulli mechanism, the "Transformed Taxpayer Identification Number," used in selecting other IRS Business Master File samples, such as for the Corporations and Partnership Studies (Harte, 1986). This permanent random number procedure was meant to improve the year-to-year estimates of change by increasing the likelihood of an entity being included in the sample in succeeding years. Clearly, this is not an issue for Estates, but it did reduce the programming burden.

The selection probabilities were set within strata, with those records with a Transformed Taxpayer Identification Number below the designated probability selected for the sample.

In addition to that selection process, a 1-percent Continuous Work History Sample (CWHS) set of ending digits for the Social Security numbers was employed. We felt that, since some of the CWHS digits were in use for the Statistics of Income Individual Study, this might allow a greater overlap between the two studies.

▶ Strata Boundaries

There are two sets of boundaries that need to be determined: age, and size of Gross Estate. Fortunately, in the later case, our task was simplified by the administrative systems. Each return was assigned a Gross Estate Code, manually, based on the size of the Estate. At the time this design was first implemented, the value itself was not available.

Gross Estate Codes, shown in Figure 3 below, with a value of less than 6 were for returns below the filing threshold in 1982, and thus were not subjected to the Bernoulli sampling. These smaller estates were filing for the record only, though we did sample them using the CWHS digits.

Figure 3.--Defining the Gross Estate Code

Size of Gross Estate	Code
Under $300,000	1 - 5
$300,000 under $500,000	6
$500,000 under $1,000,000	7
$1,000,000 under $5,000,000	8
$5,000,000 or More	9

Determining the age groups was a more difficult problem. The sample has to address two populations: the estates affected by the tax law and the living population for the Personal Wealth Estimates. In addition, we made the assumption that the age distributions within the Gross Estate categories would have a significant impact; so, we planned separate age classes for the various Gross Estate Codes. The reasoning was that, as age increases, the opportunity to accumulate wealth also increases. Thus, the median age for the smaller estates' decedents would be less than that for larger estates.

The data we had available at that time were from the 1977 Estates Study, which as we noted above had but three strata based on the size of Gross Estate. The estimates were tallied into 5-year bands. As one might expect, given the nature of the population under consideration, most of the low age-groups were empty of observations.

Over the years from 1977 to 1982, though, the number of estates in each category grew, even as the total number declined due to a rise in the filing threshold. This growth resulted from both inflation effects and the normal growth of the economy.

That growth adjustment only addresses the expected filing volume, not the population of interest. To address this, we need a further adjustment to predict the population of the living wealthy. That adjustment was the inverse of the mortality rate developed by the National

Center for Health Statistics, NCHS (then, in 1980, the data were in a pamphlet; now, they are available on their Web site).

The main reason for using the estimated wealthy population instead of the expected filings of estate Tax Returns is that we wished to focus on the scarcity of "youthful decedents." This mortality- weighted set of estimates allowed us to determine, in effect, what age a "youthful decedent" might be.

We used the Dalenius-Hodges' cumulative square root of the frequency method to find reasonable strata boundaries, with a goal of choosing five groups (Dalenius and Hodges, 1959). In the end, a sixth was added because there were a fair number of cases where there was no age reported. In later years, this "Age Unknown" group was folded into the highest- age category because research showed that these decedents actually were members of that group, and the numbers became quite small.

While the strategy outlined above was applied to the estates within the focus year, some felt that, with appropriate "aging" of assets for decedents from other years, we might be able to create better Personal Wealth estimates. Hence, as is seen in Table 1, some strata are reserved for "young," nonfocus-year decedents.

The later sample design tables show this strategy was revisited after the first focus year, and the strata for nonfocus-year filings expanded, duplicating the strata outline of the focus year. This revision reflected an increase in funding for this series of projects, as well as better meeting the need for data on the annual processing operations.

▶ Sample Allocation

Weighted strata variances for the value of Gross Estate (the value of all of an estate's assets) were available from the prior 1977 study. Since the data collection is from administrative records, without any costs related to contacting a taxpayer, we simply assumed that the costs were essentially the same regardless of the stratum. The sample size was set at about 13,000 records per year.

Neyman Allocation (with a set sample size or otherwise) also requires a population estimate. Since we are primarily interested in the effect of the tax law as it is applied in a given year, and that law has effects on the living as well as the estates, the appropriate population was the same as the one used to find the age-strata breaks.

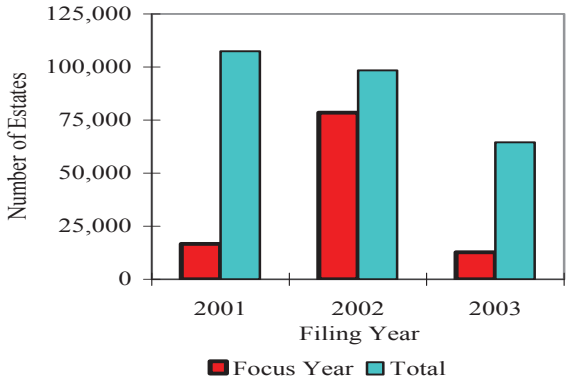
For the initial 1982 study, we allocated sample to strata under the plan for sampling the returns over 3 years, concentrating only on the year of death of the decedent, and ignoring the year of filing the administrative record.

Since the "Personal Wealth" population is more numerous than the Estates population, there were a lot of cases where the allocation prescribed more sample than there were expected estate filings. Thus, the allocation was reiterated several times, removing the certainty strata each time, before the final design's sample sizes were derived.

These sample sizes, when divided by the expected filing volumes, became the sampling probabilities used in the Bernoulli selection. These are the sampling rates shown in Tables 1 through 5, below, exclusive of the CWHS sample selections.

As a result of the filing pattern, as in the example shown in Figure 4, only about 15 percent of the sample, or about 2,000 records, were to be designated in the first year of the study, and a similar amount in the final year of the set.

Figure 4.--Estates For Decedents Who Died During 2001



Starting with the 1986 Estates Study, while the allocation of the sample to the focus year was set at the target 10,000 to 15,000 records, the difference between the expected sample size in any given filing year and the target was allocated to the nonfocus- year records within a filing year. Thus, using 2005 as an example (Focus Year 2004), while the overall sample size is about 10,000 records, about 3,000 were allocated to estates of decedents who died before 2004 or in 2005.

The allocation for nonfocus-year returns used the expected filing volume of returns, instead of the population of the wealthy used in the allocation for the focus-year strata.

▶ Changes--1986 to 2004

The initial design, in Table 1, shows the result of having age stratification dependent on the Gross Estate class. Although we show a zero probability of selection for the "Under $300,000" Gross Estate classes and other strata, those records were subjected to the 1-percent CWHS selections.

For the 1986 version of the design, shown in Table 2, the age groups were made independent of Gross Estate and were replicated for the nonfocus- year decedents. This also resulted in new age boundaries.

(Note, in this table and in subsequent ones, we will not show the classes that fall below the filing threshold due to space constraints. We used red to highlight the changes as well.)

The 1989 edition of the design, Table 3, also shows only a minor change: the introduction of an age group "65 under 75."

The next significant change arose for the 1992 study (Table 4). Here, we were finally able to replace the Gross Estate Code with the actual amount and thus expand the stratification. This design outline stood for about a decade.

The anticipated changes to the Estate Tax Law in 2001 left the design, Table 5, in some question. As a result, instead of planning to select the earliest filings for the Focus Year (2001 decedents) at the same rates as filings in later years, we planned on the initial year's sample to support estimation by itself. The focus-year pattern was also amended; so, the Statistics of Income studies will coincide with the Federal Reserve Board's Survey of Consumer Finance.

As of this writing, the tax law is still subject to change, but at least one update, having the strata boundaries match the filing thresholds, is planned for 2007.

▶ Future Research

The current trend for the tax law suggests that, in a few years, we will be canvassing the entire population, and, under some legislation, this part of the tax code would expire. However, at some future time, there may again be reason to sample a successor tax return, for one lesson from history is certainly that the Estate Tax may someday be revived. We hope that, should that arise, this paper might be of some help to that future statistician.

One more immediate issue that the Estates and Personal Wealth studies have is that the original filings on which they are based may be prone to errors in the reporting, and especially underreporting of financial assets. When such problems are discovered, the executor or lawyer will file amended returns. While such amendments are possible with other types of tax filings, because the sole person knowledgeable about the various holdings for an estate has passed away, it may be that the effect would be more serious. At this time, we simply do not have the data to examine this issue.

However, we are starting to accumulate a database that might permit such research in a few years.

▶ References

Dalenius, T. and Hodges, J.L. Jr. (1959),"Minimum Variance Stratification," Journal of the American Statistical Association.

Johnson, Barry W. and Woodburn, R. Louise (1994), "The Estate Multiplier Technique, Recent Improvements for 1989," Compendium of Federal

Estate Tax and Personal Wealth Studies, Publication 1773 (4/94), Department of the Treasury, Internal Revenue Service.

Harte, J. M. (1986), "Some Mathematical and Statistical Aspects of the Transformed Taxpayer Identification Number: A Sample Selection Tool Used at IRS," *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Kopczuk, Wojiech and Saez, Emmanuel (2004), "Top Wealth Shares in the United States, 1916-2000: Evidence From Estate Tax Returns," Working Paper 10399, National Bureau of Economic Research (http://www.nber.org/papers/w10399).

Table 1.--Strata and Selection Probabilities, 1982

Age of Decedent	*Size of Gross Estate (Based on Gross Estate Code)* *Under $300,000*	*$300,000 under $500,000*	*$500,000 under $1,000,000*	*$1,000,000 under $5,000,000*	*$5,000,000 or More*
Decedent Died in 1982					
Under 45	0	1.00	1.00	1.00	1.00
45 under 55		0.50	1.00		
55 under 60		0.35			
60 under 70			0.50		
70 or Older		0.10	0.25		
Unknown		0.10	0.25		
Decedent Died in a Year Other Than 1982					
Under 45	0	1.00	1.00	1.00	1.00
45 or Older, or Unknown	0	0	0	0	1.00

Table 2.--Strata and Selection Probabilities, 1986

Age of Decedent	*Size of Gross Estate (Based on Gross Estate Code)* $500,000 under $1,000,000	$1,000,000 under $5,000,000	$5,000,000 or More
Decedent Died in 1986			
Under 40	1.00	1.00	1.00
40 under 50	1.00	1.00	1.00
50 under 65	0.35	1.00	1.00
65 or Older, or Unknown	0.07	0.50	1.00
Decedent Died in a Year Other Than 1986			
Under 40	1.00	1.00	1.00
40 under 50	0.25	0.35	1.00
50 under 65	0.04	0.50	1.00
65 or Older, or Unknown	0.01	0.01	1.00

Table 3.--Strata and Selection Probabilities, 1989

Age of Decedent	*Size of Gross Estate (Based on Gross Estate Code)* $500,000 under $1,000,000	$1,000,000 under $5,000,000	$5,000,000 or More
Decedent Died in 1989			
Under 40	1.00	1.00	1.00
40 under 50	1.00	1.00	1.00
50 under 65	0.50	1.00	1.00
65 under 75	0.12	0.50	1.00
75 or Older, or Unknown	0.12	0.50	1.00
Decedent Died in a Year Other Than 1989			
Under 40	1.00	1.00	1.00
40 under 50	0.25	0.35	1.00
50 under 65	0.05	0.06	1.00
65 under 75	0.03	0.05	1.00
75 or Older, or Unknown	0.03	0.05	1.00

Table 4.--Strata and Selection Probabilities, 1992

Age of Decedent	Size of Gross Estate: $600,000 under $1,000,000	$1,000,000 under $2,000,000	$2,000,000 under $5,000,000	$5,000,000 under $10,000,000	$10,000,000 or More
Decedent Died in 1992					
Under 40	1.00	1.00	1.00	1.00	1.00
40 under 50	1.00	1.00	1.00	1.00	1.00
50 under 65	0.22	0.44	1.00	1.00	1.00
65 under 75	0.10	0.20	0.40	1.00	1.00
75 or Older, or Unknown	0.03	0.06	0.18	1.00	1.00
Decedent Died in a Year Other Than 1992					
Under 40	1.00	1.00	1.00	1.00	1.00
40 under 50	0.15	0.20	1.00	1.00	1.00
50 under 65	0.06	0.11	0.33	1.00	1.00
65 under 75	0.06	0.11	0.33	0.45	1.00
75 or Older, or Unknown	0.03	0.05	0.16	0.22	1.00

Table 5.--Strata and Selection Probabilities, 2001

Size of Gross Estate	Age of Decedent: Under 40	40 under 50	50 under 65	65 or Older
Decedent Died in 2001				
$675,000 Under $1,000,000	1.00	1.00	1.00	0.13
$1,000,000 under $1,500,000	1.00	1.00	1.00	0.20
$1,500,000 under $2,000,000	1.00	1.00	1.00	0.20
$2,000,000 under $3,000,000	1.00	1.00	1.00	0.40
$3,000,000 under $5,000,000	1.00	1.00	1.00	0.80
$5,000,000 under $10,000,000	1.00	1.00	1.00	1.00
$10,000,000 or More	1.00	1.00	1.00	1.00
Decedent Died in a Year Other Than 2001				
Under $1,000,000	1.00	0.01	0.01	0.01
$1,000,000 under $1,500,000	1.00	0.01	0.01	0.01
$1,500,000 under $2,000,000	1.00	0.01	0.01	0.01
$2,000,000 under $3,000,000	1.00	0.02	0.02	0.02
$3,000,000 under $5,000,000	1.00	0.04	0.04	0.04
$5,000,000 under $10,000,000	1.00	0.11	0.11	0.11
$10,000,000 or More	1.00	1.00	1.00	1.00

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.

UPDATING TECHNIQUES FOR ESTIMATING WEALTH FROM FEDERAL ESTATE TAX RETURNS

Barry W. Johnson, Internal Revenue Service
Statistics of Income OP:RS:S:SS:S, P.O. Box 2608, Washington, DC 20013-2608

Key words: Sampling probability; mortality rates; weight trimming; variance; wealth estimates

The Statistics of Income Division (SOI) of the IRS, utilizing an estimation methodology first introduced ninety years ago in Great Britain, produces estimates of personal wealth for an important segment of the U.S. population from tax returns filed for wealthy decedents. Federal estate tax returns provide a rich source of financial and demographic data on the nation's wealthiest individuals. Using data from a sample of these returns to produce wealth estimates for the living population provides a unique opportunity to study, in detail, the characteristics of the most influential individuals in the United States. In this paper, I will focus on the design used to select a sample of estate tax returns and weighting techniques used to produce estimates of personal wealth for 1992 and 1995. Weights are derived from SOI sample weights, national mortality rates, and a factor reflecting the fact that the wealthy live longer than the general population. Weights at the extreme ends of the distribution are constrained, and other methods are used to reduce the sampling variance.

Background

The first estimates of national wealth produced using death records date to the middle of the 19th century. However, early European practitioners tended to focus on developing a single weight that was applied to national totals. British Statistician Bernard Mallet [1908] was the first to use age-specific mortality rates to produce national estimates. In his 1908 estimates of wealth for 1905 and 1906, he created multipliers, within age categories, using national mortality tables and applied these to data from British Estate Duty records. Similar estimates were first produced for the U.S by Horst Mendershausen (1922-40) and later by Robert Lampman [1962] and James Smith [1994]. The Statistics of Income (SOI) Division has been using the estate multiplier technique to estimate the wealth of living individuals since 1962 (see Scheuren, 1994).

The personal wealth estimates presented in this article are based on data from Federal estate tax returns – Form 706. A decedent's estate has up to 9 months to file an estate tax return, and use of a 6-month extension is not uncommon. It is, therefore, necessary to sample returns filed over a number of calendar years in order to capture data representative of all estate tax decedents dying in a single year. In the recent past, SOI has combined returns filed over a 3-year period to produce estimates of wealth for any particular year. The estimates presented here for 1992 continue this practice. The preliminary estimates for 1995, however, are based on 2 filing years, adjusted for the remaining, unfiled returns. This was done in an attempt to provide more timely estimates; updated 1995 estimates will be published in the future. One of the strengths of the estate multiplier technique is the large sample upon which the estimates are based. The 1992 sample includes nearly 16,000 returns; the 1995 sample is made up of over 15,000 returns. Both samples are considerably larger than samples selected for other studies at comparable levels of wealth.

While the sample size and richness of available data make this estimation technique attractive, there are limitations that must be recognized. The most important is that estate tax returns provide a presumably random sample, stratified by age, not of the total population, but of living persons with gross assets at or above the filing threshold, which was $600,000 for the period of these estimates [Lampman, 1962]. Research has proven that individuals who are economically or socially better off live longer and are healthier than the general population. Factors such as access to better health services, better diet and nutrition, fewer risks on the job, and access to better housing all seem to contribute to this phenomenon [Menchik, 1991]. Therefore, it is important to determine a mortality rate appropriate to this sample. If mortality and wealth are correlated then biased estimates will result using mortality rates unadjusted for wealth level. Evidence suggests that there is an inverse relationship between these factors meaning that the multipliers will be too low and thus undervalue wealth [Smith, 1994, p. 336]. Further, it has been shown that, while patterns of wealth holding appear quite robust over a variety of reasonable alternate assumptions about the magnitude of the multipliers, overall aggregate estimates are relatively sensitive to the selection of the mortality rates. This suggests that care should be taken not to give wealth concentration estimates undue emphasis [Scheuren, 1994, p. 358].

Estate Study Sample Design

The SOI Estate Study runs on a 3-year cycle. The sample is designed mainly to accommodate year-of-death estimates, with each study concentrating on decedents dying in the first year, the focus year, of the 3-year cycle. However, the sample is adequate for filing-year estimates as well. Year-of-death estimates

are desirable because filing extensions and other filing delays mean that returns filed in any given calendar year can represent decedents who died in many different years. This means that the estate tax return data for a filing-year can reflect different economic and tax law conditions. By concentrating on a single year of death, these limitations can be overcome, making it possible to study the data in the context of a single time period.

The sample for the Estate Study is a stratified random sample with three stratifying variables. Since 1982, the stratifying variables have been year of death (focus year, nonfocus year), total gross estate, and age at death. Gross estate is divided into 5 categories: \$600,000 < \$1 million, \$1 million < \$2.5 million, \$2.5 million < \$5 million, \$5 million < \$10 million, and \$10 million or more. Age at death is divided into age < 40, 40 < 50, 40 < 65, 65 < 75, and 75 and older. Sample rates vary from 3 percent to 100 percent, with over half the strata selected with certainty. Returns are selected for the sample as they are processed for revenue purposes.

Weights for the estate sample are calculated in several steps. The first step is to adjust population and sample counts for returns that were selected into the sample but that, upon close examination, did not conform to SOI standards, or because the return data did not fall within the parameters of the study. This occurs mainly when a return is not complete by the filing deadline. In such cases, a final return will be filed when all the required information has been compiled. There are also a small number of returns that are unavailable to SOI because they are under review by other areas of the IRS. Next, adjustments are made for misclassified returns, which arise primarily from taxpayer, or IRS processing, errors that result in returns being assigned to an incorrect strata at the time sampling took place. Finally, the data are poststratified, using auxiliary data from the IRS masterfile that have been examined and corrected in an attempt to correct for large returns not originally available for sampling due to data transcription errors.

Although the overall sample of estate tax returns is large, the number of young (under 40 years of age) or extremely wealthy (gross assets of \$5 million or more) decedents tends to vary from year to year and is relatively small in comparison to their representation in the living population. The limited number of returns filed each year for decedents who were young or very wealthy can make results for these categories subject to considerable variance [Smith, 1994, p. 335]. This may create significant short-term fluctuations in the estimates attributable solely to the 'sample variance' associated with these two groups. To lessen the effect of these variations, the sample is 'smoothed' by including all returns for individuals with these characteristics filed between 1992 and 1994 (for 1992 estimates) and 1995-1996 (for 1995 estimates), without regard to the year-of-death. These segments of the sample are then poststratified and re-weighted to represent the true decedent populations in 1992 and 1995, respectively. This technique reduces the effect of outliers on estimates of the type and amount of wealth held by the young and very wealthy.

Adjustments for Missing Returns

One of the main objectives of the 3-year estate study sample design is to compute year-of-death estimates for the focus year of death. In general, most returns for year-of-death Y are filed in year Y+1. However, there are a number of returns that are filed after year Y+2 when the 3-year cycle is completed. For this reason, an adjustment, similar to a nonresponse adjustment, is computed for the focus year-of-death to account for those returns filed after year Y+2. The same type of adjustment is then computed so that year-of-death estimates can be computed using just the first 2 years in the 3-year study, allowing for more timely estimates.

Estate tax data collected by SOI for returns filed for the period 1986-1995 were used to compute the nonresponse adjustments. They were then validated using data from the IRS masterfile for the same calendar years. The adjustments were computed by first estimating the total population of filers for several years-of-death from the SOI data. These estimates were compared to population estimates based on 2 and 3 filing years and ratios then computed. The ratios were calculated using the original stratifying variables, age and size of gross estate, as well as a variable indicating whether or not the estate incurred a tax liability. Tax status was considered since it seemed likely that returns for estates incurring a significant tax liability might take longer to prepare than those for estates that, for a variety of reasons, would not incur any tax liability. The resulting ratios were then tested against both the RTF file and estimates using 2 and 3 year files with the SOI samples. Tax status was an important factor in determining when a return was filed. In addition, separate adjustments were necessary for estates with over \$10 million in gross assets. For 3-year files, ratio adjustments based on a decedent's age and tax status best approximated the estimates of the 'true' population totals for each focus year examined. For 2-year files, the adjustments based on the size of gross estate at death and tax status performed the best. The final ratios are given in Tables 1 and 2.

Mortality Differentials

One of Bernard Mallet's colleagues criticized using the national mortality rates for the wealth estimates. He suggested that mortality rates for

Table 1: 2-Year Ratio Adjustments

Adjustment Cell	Adjustment Ratio
Age < 40	
Taxable	1.18777
Nontaxable	1.26316
All age >= 40, Taxable	
$600,000 <= TGE < $1 Million	1.06129
$1 Million <= TGE <= $5 Million	1.08177
$5 Million <= TGE <= $10 Million	1.12023
TGE >= $10 Million	1.14074
All age >= 40, Nontaxable	
$600,000 <= TGE < $1 Million	1.10993
$1 Million <= TGE <= $5 Million	1.15853
$5 Million <= TGE <= $10 Million	1.23245
TGE >= $10 Million	1.22710

Table 2: 3-Year Ratio Adjustments

Adjustment Cell	Adjustment Ratio
Gross Estate > $10 Million, all ages	
Taxable	1.00178
Nontaxable	1.01414
Gross Estate < $10 Million, Taxable	
Age < 40	1.02443
40 <= Age <= 50	1.02061
50 <= Age <= 65	1.02281
65 <= Age <= 75	1.00753
Age >= 75	1.00543
Gross Estate < $10 Million, Nontaxable	
Age < 40	1.06146
40 <= Age <= 50	1.04868
50 <= Age <= 65	1.03069
65 <= Age <= 75	1.01877
Age >= 75	1.01629

"families of the peerage," or mortality tables derived from life insurance data would be more appropriate. There have been a considerable number of attempts to quantify mortality differences between the general population and the wealthy, looking at factors such as education, income, and occupation, but focusing mainly on white males. In fact, very little research has focused on the effects of these factors on the mortality of women. The first U.S. estimates of personal wealth from estate tax returns used mortality data supplied by the Metropolitan Life Insurance Company for large, whole life insurance policies to compute an adjustment factor that was then applied to the overall U.S. mortality rates. Similar data have been used by SOI for previous estimates. One drawback to this practice has been the inability to calculate sex-specific differentials from this data. Thus, an alternate data set, the National Longitudinal Mortality Study (NLMS), produced by the National Institutes of Health, is explored here.

The NLMS is a random sample of 1.3 million Americans of all ages, races, and sexes, in the civilian, noninstitutionalized population. The sample was drawn mainly from the Census Bureau's Current Population Survey. Interviews, done by telephone, had a 96-percent response rate. Respondents were at least 14 years of age. Mortality was determined by linking the Census data to the National Death Index.

Because the NLMS did not contain information on a respondent's wealth, income and occupation were used to compute the mortality differentials. Using occupation data coded from a sample of Federal estate tax returns, it was determined that a majority of decedents, for whom an occupation was reported, were employed as professionals, managers, sales persons, or farm owners/managers; the computation was, therefore, limited to NLMS respondents in those occupation categories. Income on the NLMS public-use file is categorized in 7 categories, with $50,000 or more as the top level. A preliminary file linking 706 decedents' data with income tax returns filed prior to death was used to choose appropriate levels of income for this analysis. Differentials were calculated within age and sex groups by comparing the mortality of all file decedents with those whose incomes and occupational characteristics were most similar to those of the estate tax decedents. The resulting mortality rate differentials are shown in Figures 1 and 2.

Figure 1: Mortality Experience of Males, U.S. National Longitudinal Mortality Study

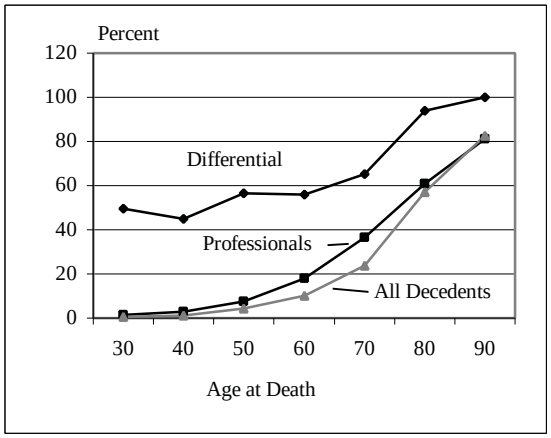
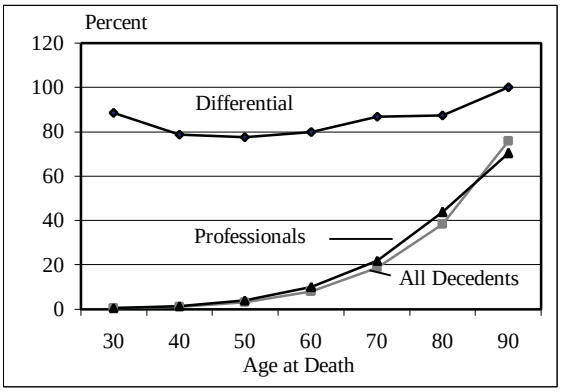


Figure 2: Mortality Experience of Females, U.S. National Longitudinal Mortality Study



The differences between the mortality of the general population and the mortality of individuals with characteristics similar to the estate tax decedent population are most pronounced for young decedents; these differences disappear entirely by age 85. Separate differentials for females were calculated for the first time and are notably smaller than those for males. The mortality differentials calculated for males are slightly larger than those derived from life insurance data, perhaps reflecting the dampening effect of the female differentials when using the aggregated life insurance data to estimate a single set of differentials for both sexes. The estimates for males seem to be in line with estimates by other researchers [see Menchik, 1991 or Wolfson Et al., 1990]. The results for both sexes are consistent with those published by the National Institutes of Health.

Multipliers

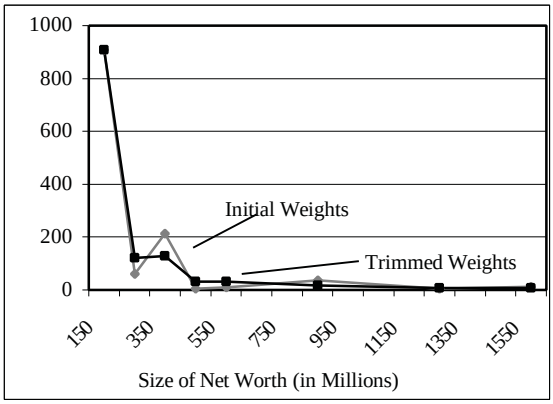
The final multipliers are calculated as:

$$MULT = \frac{\text{estate sample weight} * \text{nonresponse adjustment}}{\text{national mortality rate} * \text{mortality differential}}$$

The multipliers used in these estimates range between 1.8 and 1876.8 for the 1992 estimates and between 2.8 and 1660.8 for the 1995 estimates. The extremely skewed distribution of net worth is of particular interest to researchers. Because the underlying sample of estate tax returns was stratified by size of gross assets, which is not highly correlated with net worth, it would be appropriate to poststratify. However, the necessary control totals are not readily available. Thus, the 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 [Atkinson, 1975, p. 300-301].

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, the empirical distribution of net worth implied by the individuals in the Forbes 400 for the years 1982-1992 was examined. The data approximated a Pareto with $\alpha = 1/2$. The SOI data for 1992 were then divided into the following net worth categories: $250 to $350 million, $350 to $550 million, and greater than $550 million. The estimate of 47 in the unbounded strata was preserved, with each case assigned the mean value for the multiplier. The multiplier values in the 2 bounded net worth categories were then fit to a pareto with $\alpha = ½$, with each case assigned the mean value. The effects of these adjustments on the distribution are shown in Figure 3.

Figure 3: Preliminary and Final Distribution of Wealth for Individuals with High Net Worth.



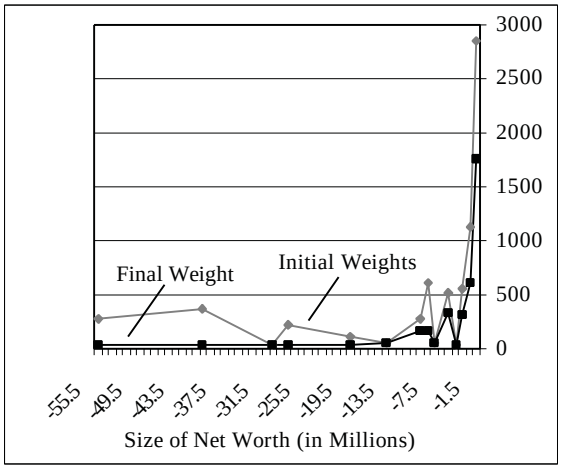
Similar adjustments were made for returns with extreme negative net worth (less than -$1 million). These cases were grouped into three categories: -$1 to -$5 million, -$5 to -$15 million, and less than -$15 million. A univariate distribution of the multipliers was computed and the multipliers trimmed at the third quartile in each of the bounded categories. There were three cases in the unbounded category. Two of these had quite large multipliers that seem unrepresentative of the general population. It was decided to assign all three cases the value of the lowest multiplier. The effects of these adjustments on the net worth distribution are shown in Figure 4.

Future Plans

Although much progress has been made since Mallot first estimated national wealth using estate duty records, several important areas for research remain. First, there is some wealth that, while not reported on Federal estate tax returns, constitutes a significant source of income for many. Life estates or income interests in assets held by a trust and defined benefit pension plans are two important income sources that are not represented in these estimates [Lampman,

1962]. As individuals shift to defined-contribution pension plans, such as 401K plans, the value of these missing assets will diminish. Even so, there would remain a significant portion of national wealth held in trusts to be explored.

Figure 4: Preliminary and Final Distribution of Wealth for Individuals with Large Negative Net Worth.



Second, although estate tax returns are generally prepared by professionals and are, therefore, likely to be more accurate in detail than survey responses, the values reported on administrative records are likely to be somewhat downwardly biased, given that they are used for the purpose of assessing taxes. This is especially true for hard-to-value assets, such as businesses and certain types of real estate. It should also be noted that the estate tax data collected by SOI are all preaudit figures. Estimates based on the results of studies of IRS estate tax return audits suggest that undervaluation may approach 5 percent of total assets, including 30 percent or more when valuing ownership interests of less than 50 percent in small companies or partnerships [McCubbin, 1994]. A nearly completed study of audit results will give us some insight into the scope and magnitude of valuation changes that result from audits. It may be possible to build in an adjustment to compensate for this bias.

Third, the wealth of individuals near death is likely to differ somewhat from that of the general population. For some, wealth will be reduced through expenses related to a final illness, while others will have made "property arrangements in anticipation of death or in recognition that an active life is over," [Lampman, 1962]. In an attempt to address this concern, data may be collected on the cause of a decedent's death. This would allow for comparisons between the portfolios of those who die suddenly and those who have planned for death carefully.

Finally, estimates of wealth derived from estate tax records are limited by the estate tax filing threshold. This limitation will be exacerbated over the next few years as that threshold rises to $1 million. However, it may be possible to extend the coverage of these estimates if comparable data can be used to estimate the wealth of individuals with gross assets under the filing threshold.

Bibliography

Atkinson, A. B., [1975] "The Distribution of Wealth in Britain in the 1960's – the Estate Duty Method Reexamined." In J. D. Smith (Ed.), *The Personal Distribution of Income and Wealth*, Columbia University Press, New York, NY, pp. 277-319.

Johnson, B.W. and Woodburn, R.L., [1994] "The Estate Multiplier Technique, Recent Improvements for 1989," *Compendium of Federal Estate Tax Data and Personal Wealth Studies*, Dept. of Treasury, IRS Pub. 1773, pp. 391-400.

Lampman, R., [1962] *The Share of Top Wealth Holders in National Wealth*, Princeton University Press, Princeton, NJ.

Mallet, B., [1908] "A Method of Estimating Capital Wealth from the Estate Duty Statistics," *Journal of the Royal Statistical Society*, Vol. LXXI, Part I, pp. 65-92.

McCubbin, J., [1994] "Improving Estimates Derived From Estate Tax Data," *Compendium of Federal Estate Tax Data and Personal Wealth Studies*, Dept. of Treasury, IRS Pub. 1773, pp. 363-370.

Menchik, P., "Economic Status as a Determinant of Mortality among Nonwhite and White Older Males: or, Does Poverty Kill?" Institute for Research on Poverty, Discussion Paper Number 93691, 1991.

_____, and Jianakoplos, N.A., [1992] "Discussion, The Estate Multiplier Technique, Recent Improvements for 1989," unpublished comments, 1992 ASA meetings.

Scheuren, F., [1994] "Historical Perspectives on IRS Wealth Estimates With a View to Improvements," *Compendium of Federal Estate Tax Data and Personal Wealth Studies*, Dept. of Treasury, IRS Pub. 1773, pp. 355-361.

Scheuren, F., [1994] "Personal Wealth, 1962," *Compendium of Federal Estate Tax Data and Personal Wealth Studies*, Dept. of Treasury, IRS Pub. 1773, pp. 153-201.

Smith, J., [1994] "Estimating the Wealth of Top Wealth-Holders from Estate Tax Returns," *Compendium of Federal Estate Tax Data and Personal Wealth Studies*, Dept. of Treasury, IRS Pub. 1773, pp. 335-354.

Wolfson, M.; Rowe, G.; Gentleman, J.; and Tomiak, M., "Earnings and Death – Effects Over a Quarter Century," unpublished manuscript, 1990.

The Author wishes to acknowledge the important contributions of R. Louise Woodburn to the research presented in this paper. Thanks are also owed to Dan Skelly, Thomas Petska, Michael Alexander, Fritz Scheuren, Martha Eller, Beth Kilss, and James Dalton. This paper reflects the views of the author alone.

The Effect of Content Errors on Bias and Nonsampling Variance in Estimates Derived From Samples of Administrative Records

Barry W. Johnson and Darien B. Jacobson
Barry W. Johnson, Statistics of Income RAS:S:SS, P.O. Box 2608, Washington, DC 20013-2608

Key words: Bias, Non-sampling error

The Statistics of Income Division (SOI) of the Internal Revenue Service (IRS) uses a number of methods for ensuring the quality and integrity of the data it produces for tax administration research. As a first line of quality assurance, codes and mathematically related data items are extensively tested as SOI employees enter them into computer databases. In addition, for a sub-sample of returns selected and processed in most studies, SOI assigns a second employee to reenter and edit the data. Values from the first and second edit are then computer-matched. A supervisor resolves discrepancies discovered during the match. The original value, second value, and correct values are all collected as a part of the quality review system, as are a set of codes that describe the cause of the error, in broad categories.

This paper will use quality review data from Federal estate tax returns (Form 706) selected into the Calendar Year 2002 SOI Estate Tax Study to estimate the effects of non-sampling error on estimates derived from the final data file.

Background

The Federal estate tax is levied on estates for the right to transfer assets from a decedent's estate to its beneficiaries; it is not an inheritance tax. A Federal estate tax return must be filed for every U.S. decedent whose gross estate, valued on the date of death, combined with certain lifetime gifts made by the decedent, equals or exceeds the filing threshold applicable for the decedent's year of death. A decedent's estate must file a return within 9 months of a decedent's death, but a 6-month extension is usually granted.

All of a decedent's assets, as well as the decedent's share of jointly owned and community property assets, are included in the gross estate for tax purposes and reported on Form 706. Also reported are most life insurance proceeds, property over which the decedent possessed a general power of appointment, and certain transfers made during life.

Expenses and losses incurred in the administration of the estate, funeral costs, and the decedent's debts are allowed as deductions against the estate for the purpose of calculating the tax liability. A deduction is allowed for the full value of bequests to the surviving spouse. Bequests to qualified charities are also fully deductible.

Data Description

The 2002 SOI Estate Tax Study was a stratified, random sample of returns filed in Calendar Year 2002 and was the second year in a 3-year study of Federal estate tax returns filed 2001-2003. The sample was designed for use in both estimating tax revenues in all 3 calendar years and personal wealth holdings for 2001 decedents. The 3-year sample period was devised to ensure that nearly all returns filed for 2001 decedents would be subjected to sampling, since a return could be filed up to 15 months after the decedent's death. The design had three stratification variables: size of total gross estate plus the value of most taxable gifts made during the decedent's life, age at death, and year of death. The year-of-death variable was separated into two categories, 2001 year of death and non-2001 year of death, in order to facilitate studies of 2001 decedents. Returns were chosen before audit examination and selected using a stratified random probability sampling method. A portion of the sample was selected because the ending digits of the decedents' Social Security Numbers (SSN) corresponded with those in the 1-percent Social Security Administration Continuous Work History Sample. However, the majority of returns were selected on a flow basis using the Bernoulli sampling method.

The sampling mechanism was a permanent random number based on an encryption of the decedent's SSN. Sample rates were preset based on the desired sample size and an estimate of the population. Sampling rates ranged from 3 to 100 percent, with more than half of the strata selected with certainty.

Data collection for the 2002 Estate Tax Study was conducted at the IRS Cincinnati Submission Processing Center. Employees entered the data from the estate tax return into a database using a Graphical User Interface (GUI) data entry system. Nearly 100 distinct data items were captured, with some balance sheet assets recurring hundreds, even thousands, of times, as assets were allocated to 32 different categories, such as stocks, bonds, and real estate. Tax returns ranged in size from a dozen to many thousands of pages, including appraisals, investment account listings, and legal documents. Tests embedded in the data entry system were used to validate entries and to ensure that mathematical relationships among variables were correctly preserved. There were more than 200 validation tests performed on each tax return included in the 2002 study.

While embedded testing can assure that codes are correct within a given range of values and that fields are mathematically consistent, many of the decisions that employees make when transforming tax return information into statistically usable data are not easily tested. For example, while several codes may be valid, determining the best code to describe a particular taxpayer's behavior or characteristics cannot always be automated. To address this problem, SOI developed a double entry quality review system. This system is a valuable tool for measuring both individual employee performance and overall data quality.

Quality Review System

A subsample of returns in the 2002 Estate Tax Study was subjected to additional review for quality assurance purposes. Returns were included in the quality review (QR) subsample through two different mechanisms, 100-percent review and product review. The 100-percent review consisted of all returns that were edited while an employee was in training. Product review was selected after the training period had been completed, and it comprised a 10-percent random sample of each employee's work. The product review sample was selected on a flow basis method using a pseudorandom number called the Transform Taxpayer Identification Number, or TTIN. The TTIN is a unique random number that is generated by mathematically transforming selected digits of the decedent's Social Security Number. The TTIN was then compared to the sample number, which represented the sample rate, in this case 10 percent. If the TTIN was less than the sample number, then the return was selected for product review.

Under the double-entry quality review system, one return was entered into the computer system twice by two different employees. The first employee did not know that a return was selected for review until after the first edit was complete, and the second employee was not allowed to see the first employee's entries. Therefore, each return had two versions in the database, the first edit and the second edit, and each was entered independently of the other.

When both employees finished editing a return, the computer compared the values from the original and QR versions. In some cases, the two versions matched perfectly; so, the return was released from the system, and the first edit data was treated as final and stored for later analysis. However, if mismatches between the two versions occurred, the discrepancies were stored in a separate data table to be reviewed by a supervisor.

The supervisor reviewed the discrepancies and charged the errors, assigning two codes to each discrepancy--one to identify the incorrect value and the other to describe the cause of the error. A discrepancy code was assigned to the error to explain which version was considered incorrect. Discrepancy codes were assigned to one of the following: the first version, the second version, both versions, or neither version. An error was assigned to both versions if both of the employees entered or interpreted the information from the return incorrectly. In this case, the supervisor was also required to supply the correct data value. In some cases an error was not assigned to either version, usually when the discrepancy was the result of a data processing peculiarity and not a true database error. After the error was assigned a discrepancy code, a numeric error resolution code was assigned to describe why the entry was incorrect. Error resolution codes indicate situations such as spelling errors, incorrect money amounts, or incorrectly assigned codes.

Once the supervisor reviewed all the discrepancies, each employee was given a list of the discrepancies, along with the discrepancy and error resolution codes, so that any first edit errors detected during quality review could be corrected prior to considering return processing complete. The feedback from the review also enabled employees to learn from their mistakes on each return and carry this knowledge into the editing of other returns. In the end, there is a database consisting of a table that includes all the values from the second edit of the return as entered, a quality review table containing a record of each discrepancy between the first and second edits (along with codes indicating who made the error and why), and a final data table containing the correct version of the return data that will ultimately be sent to customers.

For this paper, only a portion of the quality review data was used for analysis. First, data that were collected during periods of training, 100 percent review, were excluded. Second, only errors that were charged to the first edit or to both edits, meaning that the error required a correction to the final data set, were retained. This was done because these errors are more representative of errors that remain in the roughly 90 percent of the 2002 estate tax sample that was not selected for quality review. Third, errors that reflected idiosyncrasies related to the edit process itself, and not true data errors, were eliminated.

Empirical Results

Quarterly accuracy rates for each employee who worked on the Estate Tax Study for 2002 were generated using the product review data (see Figure 1). These rates were calculated using the number of returns that had at least one error charged to the first edit divided by the total number of returns that had been selected for quality review. The accuracy rates for all of the employees are not very high. However, these rates are a return level measure; any return with one or more errors is considered incorrect. The Form 706 includes an average of 150 data entry fields, while

complex returns can have more than a thousand entries; so, the probability of making just one mistake is very high. In fact, the average number of errors for each return is only 6.3.

Traditionally, supervisors have focused quality improvement efforts on those fields that are in error most frequently. By looking at the occurrence of variables *ex-ante*, using the first edit data, and *ex-post*, using the final corrected data file, it is possible to identify the frequency of original edit errors in the quality review sample. Figure 2 shows the percent changes in frequencies for variables on the file; each diamond represents a different variable. Frequencies change because many variables on the file represent balance sheet items, assets like stocks, bonds, mutual funds, and various types of real estate, which are not necessarily present in each decedent's portfolio. When an asset is incorrectly classified, not only does it change the dollar value of estimate, it also changes the frequency of occurrence of that particular attribute or asset type in the population estimates. This can be particularly problematic if the asset is of special interest to researchers. For example, there has been much discussion in the press about providing estate tax relief to small business owners. Errors that either under- or overcount the number of estates that have small businesses could have an impact on this debate. The percentages shown on the graph represent the aggregate correct frequency in the overall quality review sample, less the aggregate number originally reported, divided by the correct number. Negative percentages indicate cases where an asset was incorrectly included on the first edit. For example, the first employee may have incorrectly classified a balance sheet entry as a publicly traded stock, while the second employee may have

Figure 1: Employee Accuracy Rates

	Accuracy Rates			
Employee	Quarter 1	Quarter 2	Quarter 3	Quarter 4
17000	46.3%	23.9%	41.7%	21.7%
17100	25.0%	0.0%	0.0%	0.0%
17200	29.2%	30.8%	31.9%	40.0%
17300	57.1%	100.0%	91.7%	33.3%
17400	52.1%	28.6%	50.0%	37.9%
17500	44.4%	24.1%	54.8%	0.0%
17600	42.2%	51.9%	33.9%	46.2%
17700	41.9%	28.6%	39.3%	34.5%
17800	49.1%	25.0%	58.5%	45.6%
17900	52.3%	34.3%	59.0%	50.0%
17001	23.1%	34.2%	18.6%	44.7%
17002	39.2%	33.3%	36.2%	45.0%
17003	22.9%	20.7%	37.8%	29.1%
17004	34.2%	31.6%	22.0%	72.7%
17005	30.8%	0.0%	0.0%	37.9%
17006	26.5%	27.7%	41.4%	42.9%

correctly classified it as a mutual fund invested in a mix of financial assets. The percent changes in frequencies are generally close to zero, but there are some notable outliers.

Figure 2: Percent Change in Frequencies, Original and Final Edits

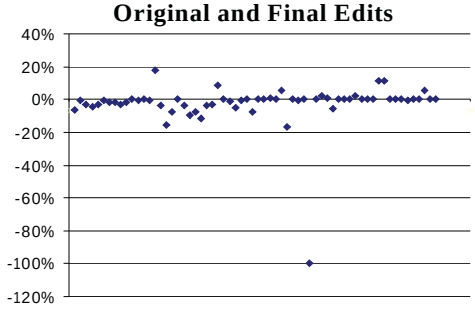
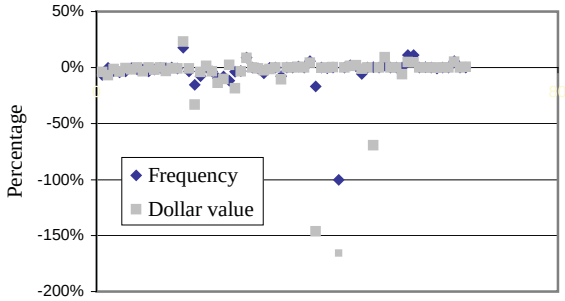


Figure 3 shows percentage changes in dollar amounts between first and second edits overlaid on the frequency differences shown in Figure 2. Each point represents a single variable on the file. While the pattern for the dollar differences is similar to that of the frequencies, with many differences close to zero, the magnitude of the dollar differences is larger for several variables. There are two variables for which the original entries resulted in aggregate dollar values that were overstated by roughly 150 percent. This highlights the potentially large effects on final estimates that can arise from even one large dollar value error, especially for variables that are not widely distributed in the overall population. Thus, it is important to monitor both the size and frequency of data entry errors.

Figure 3: Percent Change in Dollar and Frequency Values, Original and Final Edits



Unweighted error statistics are clearly useful for monitoring data quality and assessing opportunities for operational improvements during a study period. However, since the SOI study of Federal estate tax returns is based on a stratified random sample of the filing population, the effect of data entry error on final population estimates derived from this sample will vary inversely with the selection rate associated with each return. Using appropriate sample weights, it is possible to use the 10-percent QR sample to estimate the effects of data entry errors on population estimates derived from the remaining 90 percent of the returns in the final

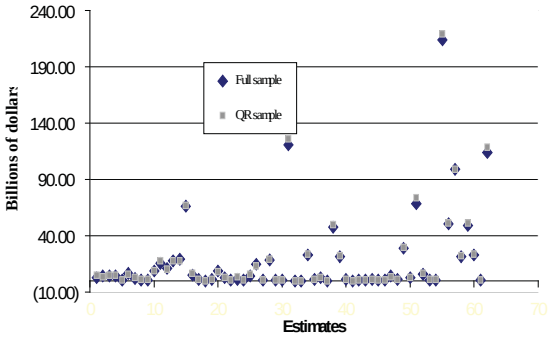
SOI data file that were not subjected to double-entry quality review. Weighted estimates provide a different perspective on the effects of nonsampling error due to the nature of the underlying estate study sample and the fact that the financial characteristics of estate tax decedents vary greatly among age and wealth classes. For example, younger decedents and those with large estates are selected into the estate tax sample with certainty and comprise more than 40 percent of the total sample file. Both groups of decedents are more likely to have had portfolios that are more complex and, thus, more subject to data entry errors than their either less wealthy, or older, cohorts. This is because many older wealth holders convert their portfolios to assets that produce tax-preferred income, usually resulting in returns that contain fewer business arrangements, which are more difficult to classify than market assets. Because the quality review sample is not stratified, weighted estimates will provide a more balanced measure of the overall effects of data entry errors on final estimates. Weighted estimates for the quality review sample were generated by using the design-based weight from the stratified estate study sample (W_s), multiplied by a quality review weight (W_q). The quality review weight itself was developed by first post-stratifying the quality review samples within the original selection strata as indicated below[1]:

Final Weight = $W_s * W_q$
Where $W_s = N_i/n_i$
Post-Stratification: $W_q = n_{if}/n_{qif}$

For some strata, the quality review sample was either zero or too small to create a post-strata cell. For these cases, strata were collapsed across age categories so that estate size classes were preserved.

Figure 4 shows full population dollar value estimates from the quality review data using the post-stratified quality review weight and compares them to population estimates using the full weighted estate study sample. Each pair of data points represents a different variable on the file. The quality review data estimates for each variable are denoted by the gray squares, and the full sample estimates are denoted by the black diamonds. For most variables, the QR sample estimates are larger than the population estimates from the full estate sample, indicating that the QR sample introduces a positive bias. This bias arises because the QR sample is a simple random sample of a stratified sample that favors large dollar value returns. In such cases, ratio raking can often be employed to decrease the bias; however, in this case, the QR sample size was insufficient in the lower gross estate size classes.

Figure 4: Full Sample vs. QR Sample Estimates



While the weighted QR data estimates are somewhat biased due to the design of the sample, they still provide an important indication of the effects of data entry errors on final estate tax sample estimates. Figure 5 shows weighted and unweighted estimates of aggregate differences between original and final values of both frequency and dollar value estimates for selected variables. A negative value means that a variable was over represented in the original, uncorrected data, and a positive value means it was originally underrepresented. Weighted results rank errors differently for some of the variables. For example, errors in classifying noncorporate business assets had a much greater impact on final weighted estimates than would have been evident had the analysis been limited to examining the unweighted QR data. Conversely, the unweighted QR data implied that the effects of errors on estimates of farm real estate were greater than they are in the final, weighted estimates. Clearly, using weighted estimates, along with the unweighted quality review data, provides a more

Figure 5: Differences between First and Final Edits

Data Element	Frequency	Dollar Value
Noncorporate Businesses	**-11.00%** *-5.29%*	**-5.79%** *-3.55%*
Closely held stock	**-3.06%** *-3.42%*	**-1.01%** *-0.71%*
Real estate	**6.70%** *6.82%*	**7.34%** *6.17%*
Farm land	**-0.91%** *-1.95%*	**-1.09%** *-3.66%*
Funeral expenses	**0.25%** *0.09%*	**0.15%** *0.04%*

Values in *italics* are unweighted estimates

[1] The subscript "if" signifies that certain reject returns were removed from the estate study sample prior to post-stratifying.

balanced method of assessing where to focus data quality improvement efforts.

Figure 6 compares the weighted percent differences between original edit estimates and final, corrected estimates with coefficients of variation (C.V.) from the full estate tax study sample in order to relate the sampling and nonsampling variances associated with selected fields. For some estimates, such as the values for noncorporate businesses and publicly traded corporations, the nonsampling error attributable to data entry is much greater than the sampling variance. For others, such as estimates of stock in closely held or untraded corporations and farm land, the sampling error, represented by the C.V., is actually greater than the nonsampling error attributable to data entry errors, indicating that data entry errors are not a significant cause of additional variance in the estimates. Fields for which nonsampling error is relatively large provide opportunities for future data quality improvement efforts.

Figure 6: Data Entry Error vs. Sample Variance

	Frequency		Money Amount	
Data Element	**% diff**	**C.V.**	**% diff**	**C.V.**
Non-corporate businesses	-11.00%	4.45%	-5.79%	3.89%
Publicly traded stock	15.02%	.78%	20.00%	1.17%
Closely held stock	-3.06%	3.47%	-1.01%	2.18%
Real estate	6.70%	1.92%	7.34%	2.19%
Farm land	-.91%	4.34%	-1.09%	4.68%
Funeral expenses	.25%	.57%	.15%	1.19%
Spousal trusts	4.25%	2.97%	1.29%	1.58%

Conclusion

There is much to be learned through careful analysis of the data generated by SOI's double-entry quality review systems. The results of these analyses can be used to improve data collection systems and enhance worker training. Information on nonsampling error should also be useful to data users who could use data quality metrics to more accurately interpret economic modeling results and to ultimately build models that are more robust.

This analysis, however, revealed that the database format and the type of data that are collected from the quality review samples make certain types of analysis difficult, if not impossible. While a complete copy of the second edit is saved for all QR returns, the original, uncorrected first edit values are not saved when first edit errors require corrections. Information on discrepancies is kept in all cases, but, because corrections can involve changing any number of related fields, it is difficult to reconstruct exactly the first employee's original entries. If more sophisticated analysis is desired, including the study of secondary errors that arise as a result of a primary data entry error, archiving a complete copy of the first edit, along with associated error reason and discrepancy codes, should be considered.

It is also important that supervisors apply error reason and discrepancy codes consistently. All too often, discrepancies are resolved by several different supervisors. Some, especially those serving in a temporary capacity, may feel a great deal of peer pressure to avoid assigning errors to individual employees, even in cases where the assignment of an error would not directly impact employee performance appraisals, such as when an error is attributable to lack of clarity in editing instructions. This inconsistency makes it difficult to measure the extent to which errors exist and to learn of ways to avoid them in the future.

Related to this problem is that the measure of employee performance currently in place is not adequate. It is simply unfair to use a return level measure of accuracy when the difficulty of the work is so variable across returns. A more balanced measure would relate the number of individual errors an employee makes to the number of fields he or she actually edited, thus giving full consideration to the number of edit decisions that were made on each return.

Finally, there are sample design issues that became apparent from this analysis. The QR sample is biased and could be improved by taking into consideration the underlying structure of the estate tax study sample design. Even this would not provide coverage of variables that are relatively rare, but perhaps important, in policy debates. To address this problem, samples could either be increased or targeted to include more returns with important characteristics, such as those filed for small business owners, or returns that, because of the types of entries made during first edit, are more likely to contain significant problems. Samples could also vary with worker skill levels. One possibility would be to develop a system that sets a weekly QR sample rate for each individual employee based on individual rolling average accuracy rates. Sample rates could be set automatically based on preset performance standards. Automating the process would avoid putting supervisors in the awkward position of having to 'punish' poor performers with additional oversight, making it easier to match feedback and training efforts to performance levels.

USING AUXILIARY INFORMATION TO ADJUST FOR NON-RESPONSE IN WEIGHTING A LINKED SAMPLE OF ADMINISTRATIVE RECORDS

Barry W. Johnson and Paul B. McMahon, Internal Revenue Service
Presented at the 2002 American Statistical Association

Federal estate tax returns are a rich source of information on the assets and liabilities associated with decedents, as well as data on beneficiaries of estates. When linked with income tax data for the decedents and their beneficiaries, the resulting data base provides a unique opportunity to study a variety of important economic issues relating to the transfer of wealth and the accumulation of capital. However, in creating such a complex, linked data base, it is inevitable that, for a variety of reasons, a number of records would be missing.

In this paper, we detail steps taken to weight the linked files. We adjust the linked record weights in two stages. First, an adjustment factor is created to balance to the original population totals, essentially treating unmatched records as non-respondents. Next, we employ auxiliary data, post-stratification, and raking to adjust the sampling weights and then compare those results to estimates from other administrative record sources.

Background

The Federal estate tax is a tax on the transfer of assets from a decedent's estate to its beneficiaries and is, therefore, levied on the estate. It is not an inheritance tax. The estate tax, the gift tax, and the generation-skipping transfer tax, together, form the Federal unified transfer tax system. This system taxes transfers made by individuals both during life and at death.

A Federal estate tax return, Form 706, must be filed for every U.S. decedent whose gross estate, valued on the date of death, combined with certain gifts made by the decedent, equals or exceeds the filing threshold applicable for the decedent's year of death. The return must be filed within 9 months of a decedent's death, unless a 6-month extension is requested and granted. All of a decedent's assets, as well as the decedent's share of jointly owned and community property assets, are included in the gross estate for tax purposes and reported on Form 706. Also reported are most life insurance proceeds, property over which the decedent possessed a general power of appointment, and certain transfers made during life.

Expenses and losses incurred in the administration of the estate, funeral costs, and the decedent's debts are allowed as deductions against the estate for the purpose of calculating the tax liability. A deduction is allowed for the full value of bequests to the surviving spouse, including bequests in which the spouse is given only a life interest, subject to certain restrictions. Bequests to qualified charities are also fully deductible.

The Statistics of Income Division (SOI) of the Internal Revenue Service selects a sample of Federal estate tax returns filed during the calendar year as part of its annual estate study. These data are used for budget analysis, tax law evaluation, and other economic studies. From time to time, a subsample of estate tax returns, collectively referred to as an "estate collation study," is selected for further processing.

The collation subsample is designed to collect additional data on decedents and the beneficiaries of their estates. Some of these data are drawn from Form 706 and supplemented with information provided in wills and trust documentation. Income tax data from Form 1040 for both decedents and beneficiaries are also linked to data from the Federal estate tax return. Bequest data, combined with income data for beneficiaries, can be used to study bequest patterns and motives (see Joulfaian, 1994), as well as to better understand the effects of inheritances on certain beneficiary behaviors (see Mikow and Berkowitz, 2000). Income tax data linked to estate tax data for decedents can be used to study such issues as the relationship between realized income and wealth (see Steuerle, 1985) and the usefulness of the life-cycle model of savings for explaining bequest behavior (see Modigliani, 1988).

The Data

The design for the 1992 Estate Collation Study had four main stages, starting with the selection of the Statistics of Income 1992 Estate Tax Return Study sample. This sample of Federal estate tax returns filed between 1992 and 1994, inclusive, was designed for use in estimating both tax revenues in all 3 calendar years and personal wealth holdings for 1992 decedents. The 3-year sample period was devised to ensure that nearly all returns filed for 1992 decedents would be subjected to sampling, given the long lag that can occur between a decedent's death and the filing of an estate

tax return, due to extensions.[1] The design had three stratification variables: size of total gross estate, age at death, and year of death. Total gross estate (the sum of all the asset valuations) was chosen as a stratifier to satisfy the first use, estimating tax revenue, and was limited to five categories:

- $600,000 under $1 million,
- $1 million under $2 million,
- $2 million under $5 million,
- $5 million under $10 million, and
- $10 million or more.

Age was selected as a stratifier, in part, because personal wealth estimation is based on death rates, which are closely correlated with age. The decedent's age at death was disaggregated into five categories: less than 40, 40 under 50, 50 under 65, 65 under 75, and 75 or older (including age unknown). The year-of-death variable was separated into two categories based on whether the year of death was 1992 or another year. This outline was designed in late 1990 and implemented in 1992, with minor sampling rate changes for non-1992 decedent strata in Calendar Years 1993 and 1994. The sampling probabilities for the 20 strata for 1992 decedent estates were not changed over the sampling period.

Estate tax returns were sampled during administrative processing, without regard to the possibility of any audit examination. A portion of the sample was selected because the decedents' Social Security number (SSN) ending digits corresponded with those in the Social Security Administration's Continuous Work History Sample (CWHS). However, the majority of returns were selected on a flow basis using a Bernoulli sampling method. The actual sampling mechanism creates a permanent random number based on an encryption of the SSN (see Harte, 1986). Sample rates were preset based on the desired sample size and an estimate of the population. They ranged from 3 percent to 100 percent, with more than half of the strata selected with certainty. These samples were limited to returns filed for decedents with total gross estates of at least $600,000, the estate tax filing threshold in effect for this period. Of the 28,530 returns sampled between 1992 and 1994, 11,943 were for decedents who died in 1992.

[1] An examination of returns filed between 1982 and 1992 revealed that almost 99 percent of all returns for decedents who die in a given year are filed by the end of the second calendar year following the year of death. Further, the decedent's age at death and the length of time between the decedent's date of death and the filing of an estate tax return are related (see Johnson, 1998). Therefore, it was possible to predict the percentage of unfiled returns, within age strata, and to adjust the final 1992 year-of-death sample weights to account for returns not filed by the end of the 3-year sampling period.

Collation Study Data

A subset of returns filed for decedents who died in 1992, and for whom an estate tax return was filed in either 1992 or 1993, was selected for inclusion in the 1992 Estate Collation Study. The subsample was limited to these 2 study years because of time restrictions for extracting the particular IRS Master File data in which we were interested. Because one study goal was to examine the relationship between income and wealth for decedents, it was necessary to have income data for, at minimum, the last full year prior to death. The source records on the Individual Master File (IMF) that we required were only retained for those posting in the current calendar year and the 2 immediately previous years (other types of records had longer retentions but contained insufficient data for our needs). Thus, in order to acquire Tax Year 1991 individual return filings (submitted in 1992), we had to cut off our selection for this collation study after the Calendar Year 1993 Form 706 selections. Estate tax returns filed during 1994 for decedents who died during 1992 had to be ignored in the sampling process. This truncation of the sample period, however, introduced significant bias since complex estate tax returns, especially those for large estates, take the most time to prepare. Much of the work documented in the rest of the paper focuses on trying to reduce the effects of this bias on estimates generated from the collation data base.

In focusing on returns filed for 1992 decedents, we eliminate 20 strata from the original estate study sample. The sample of 1992 decedents was itself further reduced from that of the original SOI sample of estate tax returns for several reasons. First, our sponsor, the Treasury Department's Office of Tax Analysis was primarily interested in the larger estates due to their expectation that only larger amounts passed to heirs would have a discernable impact on their behaviors. Second, some of the individual income tax return data were to be collected by taking advantage of the Statistics of Income Individual Program's panel selection procedures. This panel operation was an adjunct to the standard stratified Bernoulli sampling that is the mainstay of that series. There was, however, a limit on the number of SSN's that could be added to that operation due to hardware constraints. The subsample rates ranged from 4 percent to 100 percent. Returns that indicated that a decedent had made bequests to living beneficiaries, but

for which important bequest information was not reported, were rejected from the final data set.[2]

At that point, we had two sampling processes and one frame constraint that affected the sample. In addition, there was one other administrative issue that should be considered. Due to the way that SOI computer operations are planned, programmed, and tested, the sample rates are developed almost 18 months prior to implementation, based on desired sample size and filing projections that are developed using prior-year data. However, there was a recession in 1992, which diminished the value of many estates. Thus, our actual sample was smaller than expected, both for the basic estate study and the collation study. The final collation study sample contained 4,525 decedent records. These estates reported 22,000 beneficiaries, including some beneficiaries whose bequests were contingent on either the death or coming of age of other, more primary beneficiaries.

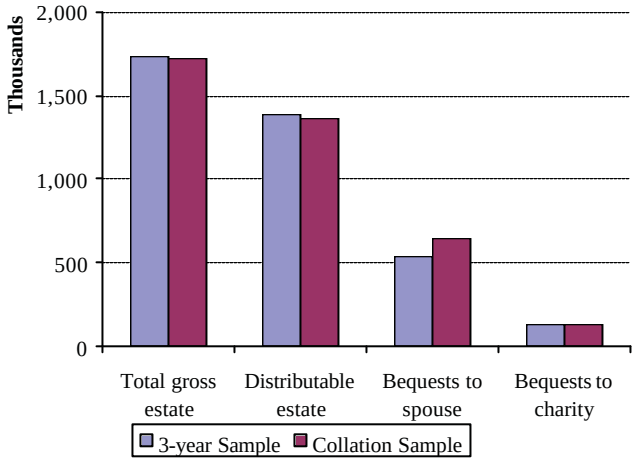
Base Weights

To calculate a base collation study weight, we needed to consider all the mechanisms that were actually involved in selecting the subsample. In order to account for the truncation of the sampling period, we post-stratified to the existing 3-year strata population counts. However, this did not fully address the reason that some returns are filed later than others. Discussions with estate tax practitioners revealed that returns reporting a significant tax liability take the longest to prepare, since several valuation experts may be consulted prior to determining final asset values, in order to minimize, as much as legally possible, the tax liability. Thus, to calculate a collation study sample weight, we further post-stratified on a binary variable indicating whether or not an estate had reported a tax liability. Note that, in both cases of post-stratification, we had the population from which the sample was drawn to tally for the strata totals. Figure A compares selected estimates using the final, weighted collation study decedent data with those from the full, weighted 3-year estate study file.

Decedent 1040 Files

For decedents in the 1992 Estate Collation Study, income tax data were obtained from the IMF for the tax period ending December 31, 1991, the last full year prior to a decedent's death. The data available were limited to those necessary for effective tax

Figure A: Mean Values for Selected Variables, 3-Year Sample vs. Collation Sample



administration. Income tax data were available for 3,767 of the 4,525 decedents in the collation sample, a linkage rate of 89.5 percent. Linkage rates varied substantially by sex and sample code. A decade ago, the IRS administrative records processing system for Individual Income Tax Returns did not include a provision for ensuring the quality of the secondary, or spousal, SSN. Since the spousal SSN on the tax form is usually that of the wife, we felt that an adjustment to the weights had to be made along gender lines to compensate for the higher level of non-matches. Indeed, while almost 90 percent of the returns filed for male decedents could be matched to a Tax Year 1991 Form 1040 return, the link rate for female decedents was only slightly more than 70 percent.

Refining this further, we found that 92 percent of the male decedents with taxable estates and 88 percent of the male, non-taxable estates were matched. This is not an important difference. Only slightly lower than those groups were the non-taxable estates of females. However, as Figure B shows, the largest difference was in the case of the taxable estates of female decedents, whose records had a match rate of only 35 percent. In fact, we were able to match only 18 percent of records for the estates of women with taxable estates valued between $2 million and $5 million. This is partly due to the very small samples in this category, which totaled only about 100 across the five size categories.

Weight adjustments for the matched 1040 returns were, thus, calculated within the original sample strata, post-stratified by gender and tax status. In several instances, samples were very small, making it necessary to collapse strata. Wherever possible, strata were

[2] In cases where a preparer had failed to provide beneficiary information on IRS Form 706, every attempt was made to collect this information from supplemental documentation, such as will and trusts. In the end, there were 22 returns that were rejected due to missing information.

collapsed across tax status, rather than sample code, to preserve the original sample probabilities. In some cases, small samples required additional aggregation. The resulting adjustments were applied to the collation study base weights. Figure C compares selected estimates using the final, weighted collation study decedent data with those from the weight-adjusted linked 1040 file.[3]

Figure B: Matching Rates For Decedents

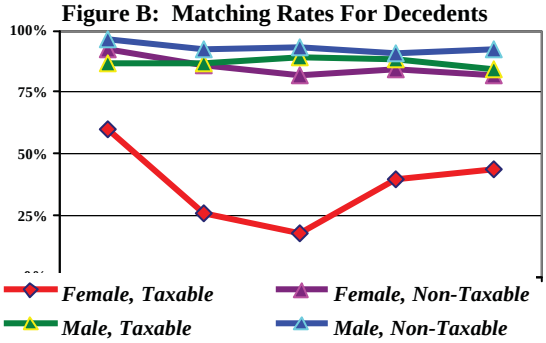
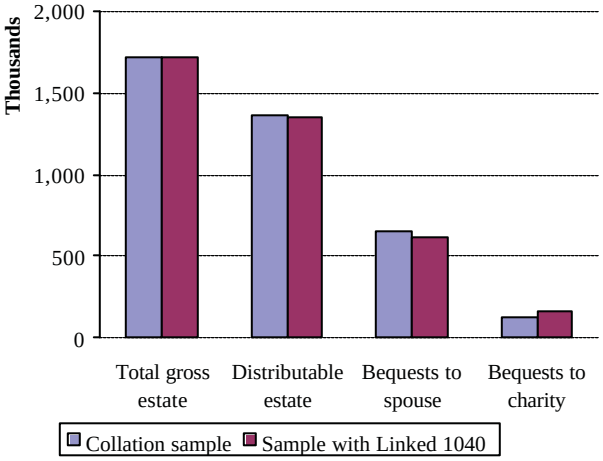


Figure C: Mean Value for Selected Variables, Collation Decedent Sample vs. 1040 Linked File



Beneficiary 1040 Files

For the purposes of studying the income effects, if any, that arise from receiving an inheritance, it was necessary to collect data from a Form 1040 filed prior to receiving an inheritance, to use as a baseline, and similar return data reflecting income after the receipt of the inheritance. For the 1992 Estate Collation Study, we selected beneficiary income tax returns for tax periods ending in 1992 (baseline) and 1995 (to see any effects of inheritance).[4]

The 1040 data for collation study beneficiaries were collected from two sources. Data for 1992 came from the IRS IMF for returns filed during Calendar Years 1992 and 1993 with tax periods ending December 31, 1992, the year of our decedents' deaths. Data for tax periods ending December 31, 1995, came from returns filed during Calendar Years 1995 and 1996 and were collected as a part of the SOI Individual Income Tax data program. Linkages were initially based on SSN matches and were confirmed by comparing name information present on Form 706 with that on Form 1040. Contingent beneficiaries (those whose inheritances were conditioned on the death, coming of age, or disclaimer of another beneficiary) were not considered in this analysis.

There were 10,983 beneficiaries for whom income tax information was available for tax periods ending 1992 and 1995, a linkage rate of 55.1 percent, much lower than that of decedents. The actual linkage rates varied substantially by sample code. An adjustment similar to that calculated for the decedent 1040 data was indicated. However, in this case, there were additional possible explanations for non-matches. First, for some beneficiaries, the preparer may have refused to provide an SSN, since it is not used for tax administration purposes. Second, for beneficiaries whose bequests were in the form of a trust, the entity identification number (EIN) associated with the trust may have been reported instead of the beneficiary's SSN. Third, transcription errors introduced either during the preparation of the original return or during data collection were also possible. Additionally, some beneficiaries may have been too young to have ever filed income tax returns in one or both periods, while others who had filed in 1992 may have died before 1995. These last possibilities introduce some uncertainty as to the exact population of beneficiaries for whom Form 1040 data should have been available. The first step in calculating final weights for this file, then, was to determine the appropriate population to use in adjusting the base weights.

In determining the population of beneficiaries whom we believed should have filed a Form 1040 in both periods, it was necessary to know the age of each beneficiary. An individual's date of birth was available

[3] These estimates differ slightly from those in Figure A because they are limited to decedents who had made bequests to living beneficiaries. A small group of decedents selected into the collation study had limited their bequests to charitable organizations.

[4] 1992 was chosen over 1991 due to the availability of more complete data for that filing year. Because of delays associated with settling an estate, beneficiaries who received inheritances from 1992 decedents would not have received them in Calendar Year 1992.

from Social Security Administration (SSA) records and was automatically present for nearly all beneficiaries for whom a Form 1040 for either 1992 or 1995 was available. For the remaining beneficiaries for whom a seemingly valid SSN had been reported, we tried linking to an SSA file, known as the Data Master One (DM1) file, which contained dates of birth. Of the 8,940 non-matched beneficiaries, we were able to obtain a DM1 file match for 2,200. Thus, age was still missing for the 5,295 beneficiaries for whom no SSN had been reported, as well as for the 1,445 beneficiaries for whom a seemingly valid TIN had been reported, but for whom no linkage to either 1040 data or the DM1 file was possible.

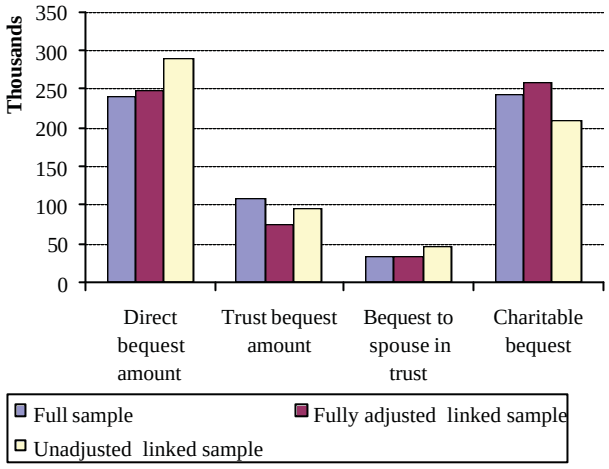
An examination of the distribution of a few key variables suggested that there was no significant difference between the groups of beneficiaries for whom age was known and those for whom age was missing. In the absence of any systematic bias, it was possible to impute missing ages using the hotdeck imputation method (see Hinkins and Scheuren, 1986). Donor cells were created, based on a beneficiary's relationship to a decedent and the decedent's age. Beneficiary age and an indicator as to whether or not a beneficiary had died prior to 1995 were selected randomly with replacement from the donor cells. Once this was completed, an examination of the data suggested that a beneficiary who was at least 18 in 1992 could have reasonably been expected to file in both periods. Consequently, beneficiaries whose actual or imputed ages were less than 18 were dropped from the analysis, as were those who had died prior to 1995. These constraints reduced the original sample of 19,926 to 18,663 non-contingent beneficiaries of 1992 estates for whom 1040 data would have been expected.

Initial weight adjustments were calculated within the original decedent sample code, thus preserving the original probabilities of selection, and were then post-stratified by tax status. The resulting initial weights were applied to the file, and weighted frequency estimates were generated by relationship to the decedent. Figure D shows that there were significant differences between the weighted estimates by relationship category for the full beneficiary sample and those produced from the linked sample. Thus, ratio raking was indicated. In addition to adjusting by relationship, we examined the possibility of separating the data further by tax status and gender. Further analysis, however, indicated that the decedent's sex was not related to the non-response bias; thus, only relationship and tax status were used. For some relationship categories, the sample was too small. So, these were combined with similar relationship categories for adjustment purposes. Adjustments were calculated and applied to the previously adjusted weights. The counts by sample code and tax status were then reproduced, using the now twice adjusted weights. Two more rounds of raking adjustments were made, each time adjusting first within the sample code and tax status and then within the collapsed relationship category. At this point, weighted frequency estimates from the matched 1040 file and the entire beneficiary file were nearly identical by both sample code and relationship category so that no more adjustments were indicated. Figure E compares selected estimates using the final, weighted collation study beneficiary data with those from the weight-adjusted linked 1040 file.

Figure D: Frequency Estimates Before Raking Adjustments

Relationship to decedent	Full sample estimate	Linked 1040 sample estimate	Percent under/over estimate
Surviving Spouse	27,023	35,274	30.5
Child	64,946	74,835	15.2
Grandchild	22,689	23,248	2.5
Sibling	11,449	9,156	-20.0
Niece/nephew	36,541	37,294	2.1
Parents	1,329	777	-41.5
Other relative	26,352	19,116	-27.5
Not related/unknown	26,953	14,823	-45.0
Total number beneficiaries	217,282	214,523	-1.3

Figure E: Mean Value, Selected Variables, Full Beneficiary File vs. Linked 1040 File



Conclusion and Future Plans

While the 1992 Estate Collation Study data base has great research potential, biases, introduced by small sample sizes and non-response problems, provide significant challenges. Particularly troubling was the necessity of truncating the sampling period from 3 to 2 years in order to conform with administrative records processing systems. Adjusting the sample weights, using post-stratification and raking, seems to be a practical method of reducing some of these biases for particular types of analyses.

The work presented in this paper suggests several additional research projects. First, the estimate for bequests through trust from the beneficiary linked data file was significantly lower than the value estimated using the full beneficiary sample file (see Figure E). This bias was not surprising, given that, while only beneficiaries with an SSN were included in the linked data file, a trust EIN was very often reported instead of an SSN when a beneficiary's entire bequest was in the form of a trust. Additional post-stratification by the form of bequest might reduce this bias. Second, we would like to measure the variances of our estimates in order to test whether differences between the mean values calculated using the linked files with adjusted weights, and those produced using the larger estate tax samples, are significant. Calculating variances, however, will require significant resources given the relative difficulty of producing variance estimates for stratified and linked datasets. Third, the post-stratification results from this work suggest that the same approach could be used to improve estimates from the annual estate study samples, although more research will be needed to determine the appropriate post-stratification classes.

Future collation studies will be affected by a number of recent developments. SOI has already undertaken a collation study of 1998 decedents with a much larger sample size. Other developments, such as IRS efforts to improve the quality of secondary SSN's on the IRS Master File and a new SOI archive of IMF data for a long time-series of tax years, should reduce some of the most troubling sources of bias present in the 1992 collation study data base. Studies beyond that of 1998 decedents will be limited by recent legislative changes that increase the estate tax filing threshold incrementally for decedents who die between 1999 and 2009 and then eliminate the tax entirely for decedents who die after December 31, 2009.[5]

[5] The Economic Growth and Tax Relief Reconciliation Act of 2001 calls for the repeal of the estate tax for decedents dying after December 31, 2009. However, that legislation expires after December 31, 2010. It is unclear, at present, whether or not the repeal of the tax will be made permanent.

References

Harte, James M. (1986), "Some Mathematical and Statistical Aspects of the Transformed Taxpayer Identification Number: A Sample Selection Tool Used at IRS," *American Statistical Association 1986 Proceedings of the Section on Survey Research Methods*, pp. 603-608.

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

Johnson, Barry W. (1998), "Updating Techniques for Estimating Wealth From Federal Estate Tax Returns," *American Statistical Association 1998 Proceedings of the Business and Economic Statistics Section*, pp. 143-147.

Joulfaian, D. (1994), "The Distribution and Division of Bequests in the U.S.: Evidence from the Collation Study, OTA Paper 71, U.S. Department of the Treasury.

Mikow, J. and Berkowitz, D. (2000), "Beyond Andrew Carnegie: Using A Linked Sample of Federal Income and Estate Tax Returns To Examine the Effects of Bequests on Beneficiary Behavior," *American Statistical Association 2000 Proceedings of the Section on Government Statistics and Section on Social Statistics*, pp. 183-188.

Modigliani, F. (1988), "The Role of Intergenerational Transfers and Lifecycle Savings in the Accumulation of Wealth," *Journal of Economic Perspectives* 2, pp. 15-40.

Steuerle, C. E. (1985), "Wealth, Realized Income, and the Measure of Well-Being," in *Horizontal Equity, Uncertainty, and Economic Well-Being*, M. David and T. Smeeding, editors, University of Chicago Press, Chicago.

The Effect of Late-Filed Returns on Population Estimates: A Comparative Analysis

by Brian Raub, Cynthia Belmonte, Paul Arnsberger, and Melissa Ludlum, Internal Revenue Service

The Statistics of Income (SOI) division of the Internal Revenue Service (IRS) collects and disseminates detailed data based on samples of administrative records, including tax and information returns. Estimates for populations of interest for SOI studies are produced by drawing stratified, random Bernoulli samples of tax and information returns as they are filed, over periods that span a predetermined timeframe. While this methodology results in the inclusion of the majority of targeted returns, a small number of returns for each study are received beyond the data collection period. These "late-filed" returns may introduce non-response bias into the population estimates, which might be mitigated by post-stratification or weighting adjustments. (The term "late-filed return" as used in this paper does not address the compliance, or lack thereof, of return filings with statutory requirements.) Using three SOI studies with varying sampling frames, this paper will function as a case study on the effects of truncated sampling periods on population estimates.

The data presented in this paper are derived from two sources—sample data produced by SOI and administrative data obtained from the IRS Masterfile for the population of returns filed. SOI sample data typically include detailed, error-perfected financial and other information about the tax filing entity. SOI sample data are used to produce population estimates that are used in statistical studies and for analysis of tax policy. Data obtained from the IRS Masterfile include limited information for the population of filers. This information is generally used for a variety of purposes related to tax administration.

SOI conducts annual studies of a wide range of filers, including individuals, corporations, partnerships, estates, trusts, tax-exempt charitable organizations, and many other filers. This paper focuses on three SOI studies—the Estate Tax study, the Private Foundation study, and the Exempt Organization study.

▶ The Estate Tax Study

With its annual Estate Tax study, SOI extracts demographic, financial, and asset data from Federal estate tax returns. The annual study allows production of a data file for each filing, or calendar, year. By focusing on a single year of death for a period of 3 filing years, the study allows production of periodic year-of-death estimates. A single year of death is examined for 3 years, as over 98 percent of all returns for decedents who die in a given year are filed by the end of the second calendar year following the year of death. Data included in this paper are for Year of Death 2004 and were obtained from returns filed in Calendar Years 2004-2006.

The estate of a decedent who, at death, owns assets valued in excess of the estate tax applicable exclusion amount, or filing threshold, must file a Federal estate tax return, Form 706, *U.S. Estate (and Generation-Skipping Transfer) Tax Return*. For decedents who died in 2004, the exclusion amount was $1.5 million. Alternate valuation may be elected only if the value of the estate, as well as the estate tax, is reduced between the date of death and the alternate date. The estate tax return is due 9 months from the date of the decedent's death, although a 6-month filing extension is allowed. In some cases, longer filing extensions may be permitted.

For the Year of Death 2004 Estate Tax study, there were 11,817 Form 706 returns in the sample selected from a population of 42,424. The SOI Estate Tax study is classified into strata based on year of death, the size of total gross estate, and age of the decedent. For the Year of Death 2004 study, there were a total of 57 sam-

This paper was originally presented at the 2009 annual meetings of the American Statistical Association held in Washington, D.C., on August 2-6, 2009.

pling strata, with sampling rates ranging from 4 percent to 100 percent.

▶ The Private Foundation and Exempt Organization Studies

The annual SOI studies of private foundations and exempt organizations collect detailed financial data, as well as information on charitable and grant-making activities and compliance with IRS regulations from information returns filed by exempt organizations. Studies are conducted for a single tax year and include samples of returns filed and processed during the 2 calendar years immediately following the target tax year. Data discussed in this paper for the Private Foundation and Exempt Organization studies were obtained for Tax Year 2004 returns filed in Calendar Years 2005 and 2006. The Tax Year 2004 samples include organizations with accounting periods beginning in Calendar Year 2004 (and ending between December 2004 and November 2005), for which returns were filed and processed to the IRS Business Masterfile during Calendar Years 2005 and 2006. While this 2-year sampling period ensures almost complete coverage of the target population, there are still a number of returns processed after the close of the second year (i.e., December 31, 2006 for the Tax Year 2004 study), which are generally excluded from the samples.

Private foundations and nonexempt charitable trusts are required to file Form 990-PF (*Return of Private Foundation or Section 4947(a)(1) Nonexempt Charitable Trust Treated as Private Foundation)* annually. Similarly, certain exempt organizations are required to file Forms 990 *(Return of Organization Exempt from Income Tax)* or Form 990-EZ *(Short Form Return of Organization Exempt from Income Tax)*. SOI conducts annual studies based on samples of Forms 990-PF, 990, and 990-EZ filed for a given tax year. These information returns are due 5 months after the close of the organization's accounting period, although a 3-month filing extension is allowed. In some cases, additional filing extensions may be granted.

For the Tax Year 2004 Private Foundation study, there were 7,805 Form 990-PF returns in the sample, selected from a population of 80,570. The SOI Private Foundation study is classified into strata based on the size of end-of-year fair market value of assets, with each stratum sampled at a different rate. Sampling rates ranged from 1 percent for private foundations with total assets less than $125,000 to 100 percent for private foundations with total assets of $10 million or more.

The Tax Year 2004 exempt organization sample of section 501(c)(3) filers comprised 15,070 Forms 990 and 990-EZ, selected from a population of 279,415. End-of-year book value of assets was the stratifying variable for the exempt organization study. Sampling rates ranged from 1 percent for exempt organizations with total assets less than $500,000, to 100 percent for those with total assets of $50 million or more.

▶ Late-Filed Returns

To examine the effect of late-filed returns on each of the studies, an augmented sampling frame, which incorporates 2 years of additional return filings, was constructed from IRS Masterfile data. The following tables show the number of late-filed returns received within the current and augmented sampling frames, as well as the percentage of selected financial variables represented by returns received inside and outside of the sampling period.

Table 1, below, shows the percentage of Year of Death 2004 Forms 706 filed, total gross estate, and net estate tax reported for returns filed over a 5-year collection period (2004–2008), by size of gross estate and by age of the decedent More than 98 percent of all Year of Death 2004. Forms 706 filed over the 5-year period were received within the 3 years, 2004 through 2006, from which returns were sampled. However, the estates of younger decedents filed returns outside of the 3-year sampling frame proportionately more often than the estates of their older counterparts. For example, nearly 4 percent of returns filed for decedents under 40 were received in 2007 and 2008. The percentage of total gross estate represented by late-filed returns was 1.1 percent, with the corresponding figure for net estate tax only 0.5 percent. These smaller percentages are attributable to the fact that late-filed returns were smaller on average

than other returns and were proportionately more often nontaxable, as shown in the following tables.

Table 1: Estate Tax Returns Filed for 2004 Decedents, IRS Masterfile Data by Age of Decedent, 2004–2008

Calendar Year	*Returns*	*Total gross estate*	*Net estate tax*
2004-2006	**98.4%**	**98.9%**	**99.5%**
Under 40	96.4%	97.0%	100.0%
40 under 50	97.0%	97.5%	98.9%
50 under 65	97.2%	97.6%	98.7%
65 and over	98.6%	99.0%	99.6%
2007-2008	**1.6%**	**1.1%**	**0.5%**
Under 40	3.6%	3.0%	0.0%
40 under 50	3.0%	2.5%	1.1%
50 under 65	2.8%	2.4%	1.3%
65 and over	1.4%	1.0%	0.4%

Table 2 examines the same population as the previous table, classified by size of total gross estate. The table shows that returns for the smallest estates, those with between $1.5 and $2 million in gross estate, were filed in the 2 years immediately following the sampling period twice as frequently as were returns for the largest estates.

Table 2: Estate Tax Returns Filed for 2004 Decedents, IRS Masterfile Data by Size of Total Gross Estate, 2004–2008

Calendar Year	*Returns*	*Total gross estate*	*Net estate tax*
2004-2006	**98.4%**	**98.9%**	**99.5%**
$1.5 million<$2.0 million	98.2%	98.2%	98.9%
$2.0 million<$3.0 million	98.3%	98.3%	99.1%
$3.0 million<$5.0 million	98.4%	98.4%	99.1%
$5.0 million<$10.0 million	98.8%	98.8%	99.5%
$10 million and over	99.1%	99.6%	99.7%
2007-2008	**1.6%**	**1.1%**	**0.5%**
$1.5 million<$2.0 million	1.8%	1.8%	1.1%
$2.0 million<$3.0 million	1.7%	1.7%	0.9%
$3.0 million<$5.0 million	1.6%	1.6%	0.9%
$5.0 million<$10.0 million	1.2%	1.2%	0.5%
$10 million and over	0.9%	0.4%	0.3%

Table 3 examines the same population as Tables 1 and 2, classified by tax status of the return. It shows that nontaxable returns were filed outside of the sampling period more than twice as often as taxable returns.

Table 3: Estate Tax Returns Filed for 2004 Decedents, IRS Masterfile Data by Tax Status, 2004–2008

Calendar Year	*Returns*	*Total gross estate*	*Net estate tax*
2004-2006	**98.4%**	**98.9%**	**99.5%**
Taxable	99.1%	99.4%	99.5%
Nontaxable	97.9%	98.2%	N/A
2007-2008	**1.6%**	**1.1%**	**0.5%**
Taxable	0.9%	0.6%	0.5%
Nontaxable	2.1%	1.8%	N/A

Table 4 illustrates the extent to which estimates based on Form 990-PF data collected from the current 2-year sampling period might be enhanced by using the augmented sampling frame. More than 98 percent of the Tax Year 2004 private foundation returns included in the augmented sampling frame were processed in the 2 years immediately following the close of the tax year. A closer examination reveals that the percentage of returns received and processed during the first 2 years increases with asset size. For example, 97.9 percent of returns filed by small organizations (those with assets less than $1,000,000) were processed during the 2005-2006 period, compared to 99.2 percent of the returns of medium-sized foundations (those with assets between $1 million and $50 million), and 99.7 percent of the returns of the largest foundations (those with assets of $50 million or more).

Table 4: Tax Year 2004 Private Foundation Information Returns, IRS Masterfile Data by Calendar Year and Size of Organization, 2005–2008

Calendar Year	*Returns*	*Assets*	*Revenue*	*Charitable Disbursements*	*Excise Tax on Net Investment Income*
2005-2006	**98.3%**	**99.5%**	**99.4%**	**99.5%**	**99.5%**
Small	97.9%	98.8%	98.7%	98.9%	98.8%
Medium	99.2%	99.3%	99.3%	99.4%	99.5%
Large	99.7%	99.6%	99.6%	99.7%	99.6%
2007-2008	**1.7%**	**0.5%**	**0.6%**	**0.5%**	**0.5%**
Small	2.1%	1.2%	1.3%	1.1%	1.2%
Medium	0.8%	0.7%	0.7%	0.6%	0.5%
Large	0.3%	0.4%	0.4%	0.3%	0.4%

Table 5 shows the breakdown of data from Forms 990 and 990-EZ returns by filing period and size of assets. As with private foundations, the vast majority of

Tax Year 2004 returns were filed in the first two years after the end the tax year. Again, a large portion of the returns filed in the final 2 years of the augmented sampling frame are from small organizations – those with total assets less than $100,000. Consequently, late filers of Forms 990 add little to the aggregate totals for most of the financial variables collected: less than 1 percent of total assets, revenue, and net worth.

Table 5: Tax Year 2004 Exempt Organization Information Returns, IRS Masterfile Data by Processing Year and Size of Organization, 2005–2008

Calendar Year	*Returns*	*Assets*	*Revenue*	*Net Worth*
2005-2006	**97.3%**	**99.3%**	**99.3%**	**99.2%**
Small	95.7%	96.4%	96.2%	96.1%
Medium	98.3%	98.8%	98.8%	97.8%
Large	99.2%	99.3%	99.3%	99.3%
2007-2008	**2.7%**	**0.7%**	**0.7%**	**0.8%**
Small	4.3%	3.6%	3.8%	3.9%
Medium	1.7%	1.2%	1.2%	2.2%
Large	0.8%	0.7%	0.7%	0.7%

▶ Current Treatment of Late Filers

Although the Estate Tax, Private Foundation, and Exempt Organization studies share a common challenge in addressing the effect of late-filed returns on population estimates, each of the three studies currently uses a different approach in dealing with this challenge.

Year of Death population estimates for the Estate Tax study include weight adjustments for late-filed returns. Weight adjustment factors are calculated based on past late filing patterns using historical data from the IRS Masterfile and are updated periodically. The aim of using these weight adjustments is to improve the overall population estimates as well as estimates for the subpopulations of returns that have historically filed late with greater frequency. As shown in Table 6, weight adjustment factors varied by size of estate, tax status of return, and age of decedent. For each size of estate and age combination, non-taxable returns received a higher adjustment factor than taxable returns. Estates with $10 million or more in gross estate received weight adjustment factors based on tax status regardless of age, as illustrated in the top portion of the table. For estates with less than $10 million in gross estate, weight adjustment factors were assigned based on tax status and age.

Table 6: Weight Adjustment Factors for Year of Death 2004 Estate Tax Population Estimates

Total gross estate ≥ $10 million		
	Taxable	*Nontaxable*
All ages	1.004	1.013
Total gross estate < $10 million		
Age	*Taxable*	*Nontaxable*
Under 40	1.036	1.052
40 under 50	1.019	1.035
50 under 65	1.018	1.028
65 and older	1.009	1.020

Table 7—shows the aggregate effect of weight adjustment factors on the Year of Death 2004 estate tax estimates. The number of returns increased about 1.5 percent compared to a 1.2 percent increase in total gross estate and less than a 1 percent increase in net estate tax. The differences in the impact of weight adjustments on these three variables is consistent with the fact that late-filed returns comprised proportionately more small returns and non-taxable returns than the population as a whole.

Table 7: Effect of Weight Adjustments on Estimates of Year of Death 2004 Estate Tax Population

[Money amounts are in millions of dollars]

	Returns	*Total Gross Estate*	*Net Estate Tax*
Unadjusted estimate	41,599	183,657	22,075
Estimate with weight adjustment	42,239	185,921	22,220
Percentage increase	1.54	1.23	0.66

In contrast, population estimates for the Private Foundation study do not include standard adjustment factors to account for returns filed after the close of the two-year sampling period. Instead, during file closeout, efforts are made to identify and include late-filed returns of private foundations that would have been sam-

pled at the 100-percent rate (i.e., organizations with fair market value of assets of $10 million or more). These include returns of organizations sampled in previous study years, as well as returns of organizations posting to the IRS Masterfile for the first time. Potentially, this can extend the 2-year sampling frame by four to five months, the typical length of time between the end of the sampling period and the creation of the final study file. Table 8, shows population estimates for selected variables from SOI's Tax Year 2004 Private Foundation study. The table includes population estimates from returns processed during the regular 2-year sampling period, as well as enhanced population estimates including adjustments for late-filed returns. Only 11 large-case, late-filed returns were added to the Tax Year 2004 sample. These returns represented 100th of 1 percent of the population estimate, and about a one-fifth of 1 percent addition to total revenue, charitable disbursements, and net investment income excise tax.

Table 8: Tax Year 2004 Private Foundation Data from SOI Estimates, Including Added Late-Filed Returns

[Money amounts are in millions of dollars. Detail may not add to totals because of rounding.]

Calendar Year	*Returns*	*Assets*	*Revenue*	*Charitable Disbursements*	*Excise Tax on Net Investment Income*
SOI two-year estimate	76,886	509,471	58,539	32,071	467
Additional data from late-filed returns	11	453	129	54	1
Enhanced SOI estimate	76,897	509,924	58,668	32,125	469
Additional data as percentage of total	0.01	0.09	0.22	0.17	0.21

The Exempt Organization study includes no weight adjustments and no attempt is made to add returns to the sample that are filed outside of the two-year sampling frame. Adjustments to the sample are made for certain organizations that file returns within the 2-year sampling period. Examples of these adjustments include rejecting short-year returns and those that file with an incorrect subsection code; and adding returns that have posted incorrectly to the Masterfile as duplicate, below the filing threshold, or with incorrect total assets.

Using IRS Masterfile data as a proxy, we can mimic the Private Foundation study's technique of processing returns from the certainty strata that are filed within five months after the close of the normal sampling period. Based on the Masterfile data, 21 large-case returns would have been added to the Tax Year 2004 sample. These returns would have accounted for a one-third of 1 percent addition to the aggregate totals for assets, revenue and net worth.

▶ Strengths/Weaknesses

These analyses reveal a number of strengths and weaknesses for each of the three approaches to the late-filer problem. The weight adjustment approach, as employed for the Estate Tax study, potentially improves the overall population estimates. It also may improve estimates for subpopulations for which returns have historically been filed late with the greatest frequency. The adjustments seem to be an effective means of counteracting any bias that may result from the existing sampling period. To the extent that late filers create bias in the Estate Tax study estimates, the weight adjustment approach may mitigate the bias.

On the other hand, the weight adjustment approach may not always be an effective method of predicting filing habits. The weight adjustments are developed from observed trends in historical data; this information may not always reliably predict future filing patterns. Although the characteristics of late filers have been relatively stable over time, significant changes to the estate tax law could alter these patterns.

The inclusion of large, late-filed returns in the Private Foundation study provides for more complete coverage of the target population by including returns that would have been selected with certainty within the defined sampling period. Additionally, this approach ensures that files are suitable for time-series analysis of a specific organization or panel of organizations. This strength may be unique to data for tax-exempt organizations, whose information returns, in most cases, are not subject to the same disclosure and confidentiality rules as data obtained from tax returns filed by other types of organizations and individuals.

The primary weakness of including large, late-filed returns only in the enhanced estimate is the inconsistency that it introduces. Slight variances in tax return processing, sample file creation or review, or sample file delivery date can affect the sampling period from which the enhanced estimate is drawn from year to year. Further, the method fails to address late-filed returns of smaller organizations, which account for the largest share of the late-filing population.

The "no-adjustment" approach that is used in the Exempt Organization study ensures a consistent sampling frame with a well-defined sampling period. This approach employs the Bernoulli sample over a 2-year period and does not include arbitrary additions or discontinuations. Because the population is framed as the estimate of filers of the 2-year period and not as the "universe" of filers, the bias does not exist.

Because, unlike the weight adjustment method used in the Estate Tax study, the "no-adjustment" approach does not attempt to account for late filers, it could consistently underestimate the number of returns filed by smaller organizations. By omitting some large case returns that are received outside of the defined sampling period, this approach also provides a somewhat less complete dataset for time-series panel analysis than does the Private Foundation study.

▶ Conclusions/Future Research

Late-filed returns present a common challenge for studies of data obtained from tax returns, such as the Estate Tax, Private Foundation, and Exempt Organization studies. Although, for each of the studies, the number of late-filed returns is modest in comparison to the number of returns filed within the defined sampling period, the absence of these returns may introduce bias into the population estimates.

Currently, each of the three studies discussed in this paper uses a unique approach to mitigate the potential bias introduced by late filers. The weight adjustment method, employed for the Estate Tax study, improves some aspects of the study's estimates, but could become distorted if filing patterns observed in historical data do not continue into the future. The enhanced Private Foundation estimate, which is obtained by including targeted returns received after the end of the sampling period, benefits time-series analysis. However, it creates inconsistencies in the year-to-year sampling period. The "no adjustment" method used for the Exempt Organization study provides a distinct sampling period, but does not address the exclusion of relatively small filers from the estimates.

The unique characteristics of late filers in each of the three studies discussed in this paper, as well as the benefits and shortfalls of using each of the three approaches to address the later-filer problem, provide a number of opportunities for further research.
This analysis will be expanded to research additional tax years and years of death in order to explore historical filing patterns. This effort will attempt to isolate an optimal sampling period that balances population coverage with timeliness of completion of the estimates. Additionally, weighting adjustments, similar to those in use for the Estate Tax study, will be developed for the Private Foundation and Exempt Organization studies. The adjustments will be examined for accuracy, as well as their effect on organization-level data from year-to-year.

Developing an Optimal Approach to Account for Late-Filed Returns in Population Estimates

Cynthia Belmonte, Brian Raub, Paul Arnsberger, Charles Day[1]
[1]IRS, Statistics of Income, 1111 Constitution Ave NW, Washington DC 20024

Abstract
Estimates for populations of interest for Statistics of Income (SOI) programs are produced by drawing stratified, random Bernoulli samples of tax and information returns as they are filed, over predetermined sampling periods that often span multiple years. While this methodology results in the inclusion of the majority of targeted returns, a small number of returns for each study are filed beyond the data collection period, potentially introducing non-response bias into the population estimates. For a given sampling period, the paper will analyze historical filing patterns to develop an approach for accounting for late-filed returns. This research will assess the weight adjustment approach currently used in SOI's estate tax study and will provide a basis for application of a similar approach in each of the exempt organizations and private foundations studies.

Key Words: non-response bias, population estimates, post-stratification, Bernoulli sampling

1. Data Sources and Background

The Statistics of Income (SOI) division of the Internal Revenue Service (IRS) collects and disseminates detailed data based on samples of administrative records, including tax and information returns. The SOI sampling frame for any given study consists of tax or information returns posted to the appropriate IRS return transactions processing system within a designated time period. Often, this time period is the statutory period within which taxpayers are required to file. For other studies, in which taxpayers may file returns over many years, sampling occurs over a designated time period in which past experience tells SOI statisticians all but a small fraction of returns will be filed. In either event, some taxpayers may file returns for the period of interest after sampling for a study has ended. Over the years, SOI has taken several approaches to adjusting for the incompleteness of its sampling frames, some on a case-by-case basis and others more uniform in nature.

Building on previous research, this paper describes three SOI studies covering tax and information returns for estates, private foundations, and exempt organizations and briefly outlines current practices for handling late-filed returns [1]. Next, the authors describe two models for predicting the proportion of late-filed Estate tax returns using several covariates. Using the 2004 year-of-death sample, the authors will then apply and evaluate the new adjustment factors by comparing results to known population totals and previous estimates derived using existing adjustment factors.

1.1 The Estate Tax Study

With its annual Estate Tax study, SOI extracts demographic, financial, and asset data from Federal estate tax returns. The annual study allows production of a data file for each filing, or calendar, year. By focusing on a single year of death for a period of 3 filing

years, the study allows production of periodic year-of-death estimates. A single year of death is examined for 3 years, as over 98 percent of all returns for decedents who die in a given year are filed by the end of the second calendar year following the year of death. Data included in this paper are for Year of Death 2004 and were obtained from returns filed in Calendar Years 2004-2006.

The estate of a decedent who, at death, owns assets valued in excess of the estate tax applicable exclusion amount, or filing threshold, must file a Federal estate tax return, *Form 706, U.S. Estate (and Generation-Skipping Transfer) Tax Return*. For decedents who died in 2004, the exclusion amount was $1.5 million. Alternate valuation may be elected only if the value of the estate, as well as the estate tax, is reduced between the date of death and the alternate date. The estate tax return is due 9 months from the date of the decedent's death, although a 6-month filing extension is allowed. In some cases, longer filing extensions may be permitted.

For the Year of Death 2004 Estate Tax study, there were 11,817 Form 706 returns in the sample selected from a population of 42,424. The SOI Estate Tax study is classified into strata based on year of death, the size of total gross estate, and age of the decedent. For the Year of Death 2004 study, there were a total of 57 sampling strata, with sampling rates ranging from 4 percent to 100 percent.

1.2 The Private Foundation and Exempt Organization Studies

The annual SOI studies of private foundations and exempt organizations collect detailed financial data, as well as information on charitable and grant-making activities and compliance with IRS regulations, from information returns filed by exempt organizations. Studies are conducted for a single tax year and include samples of returns filed and processed during the 2 calendar years immediately following the target tax year. Data discussed in this paper for the Private Foundation and Exempt Organization studies were obtained for Tax Year 2004 returns filed and processed to the IRS Business Masterfile during Calendar Years 2005 and 2006. While this 2-year sampling period ensures almost complete coverage of the target population, there are still a number of returns processed after the close of the second year (i.e., December 31, 2006 for the Tax Year 2004 study), which are generally excluded from the samples.

Private foundations and nonexempt charitable trusts are required to file Form 990-PF (*Return of Private Foundation or Section 4947(a)(1) Nonexempt Charitable Trust Treated as Private Foundation)* annually. Similarly, certain exempt organizations are required to file Forms 990 *(Return of Organization Exempt from Income Tax)* or Form 990-EZ *(Short Form Return of Organization Exempt from Income Tax)*. SOI conducts annual studies based on samples of Forms 990-PF, 990, and 990-EZ filed for a given tax year. These information returns are due 5 months after the close of the organization's accounting period, although a 3-month filing extension is allowed. In some cases, additional filing extensions may be granted.

For the Tax Year 2004 Private Foundation study, there were 7,805 Form 990-PF returns in the sample, selected from a population of 80,570. The SOI Private Foundation study is classified into strata based on the size of end-of-year fair market value of assets, with each stratum sampled at a different rate. Sampling rates ranged from 1 percent for private foundations with total assets less than $125,000 to 100 percent for private foundations with total assets of $10 million or more.

The Tax Year 2004 exempt organization sample of section 501(c)(3) filers comprised 15,070 Forms 990 and 990-EZ, selected from a population of 279,415. End-of-year book value of assets was the stratifying variable for the exempt organization study. Sampling rates ranged from 1 percent for exempt organizations with total assets less than $500,000, to 100 percent for those with total assets of $50 million or more.

2. Current Treatment of Late-Filed Returns

SOI's estate, private foundation, and exempt organization studies all share a common challenge in accounting for returns filed after the end of the designated sampling period. The Estate Tax study Year-of-Death estimates include weight adjustments for late-filed returns. Such adjustments were first developed in 1997 by Woodburn, and later updated in 2007 by Raub. Weight adjustment factors are calculated using historical data from the IRS Masterfile, and vary by size of estate, age of decedent, and tax status of return. The aim of using these weight adjustments is to improve the overall population estimates, as well as the estimates for the subpopulations of returns that have historically filed late with greater frequency. To the extent that late-filers create bias in the Estate tax estimates, this approach seems to be an effective strategy in mitigating this bias. Another strength of this approach is that the data used to calculate the adjustment factors are readily available in the IRS Masterfile.

In contrast to the estate tax study, population estimates for the private foundations study do not include standard adjustment factors to account for returns filed after the close of the 2-year sampling period. Instead, during file closeout, efforts are made to identify and include late-filed returns that would have been sampled at the 100-percent rate (i.e., organizations with fair market value of assets of $10 million or more). This allows for more complete coverage of the target population by including returns that would have been selected with certainty. This allows for time-series analysis of a specific organization (or panel of organizations). Potentially, this treatment can extend the two-year sampling period by 4 to 5 months, the typical length of time between the end of the normal sampling period (in December) and the creation of the final study file (in mid-May). This can introduce some inconsistency from year-to-year, since the slightest variation in the Master File processing cycle, file review schedule, or final delivery date can affect the sampling period from one year to the next. Additionally, this method does not specifically address smaller organizations, which account for the largest share of the late-filing population.

3. Methodology and Results

The goal of the current research is to determine whether the current estate tax study adjustment factors still accurately reflect taxpayer behavior. Additionally, the authors seek to develop and assess alternative methods of estimating adjustment factors on the estate tax study, and whether such methods can be applied to other studies (Private Foundations, Exempt Organizations) that are subject to similar late-filing challenges.

The authors propose adjusting the weights of the returns in the estate tax return sample by multiplying by the inverse of the predicted proportion of returns filed by the cutoff of sampling. In order to not overly inflate variance, it is desirable that a relatively small number of adjustments be applied to the returns. Rather than attempting to calculate an

adjustment based on each return's values of selected covariates, the adjustment factors were calculated for specific categories that are either sampling strata, groups of strata, or subsets of a stratum. Such an adjustment accounts for returns that will be filed after the end of the sampling period for the estates of decedents who died during the reference year.

Discussions with the estate tax study analyst yielded three possible explanatory covariates: size of the estate (measured by the total gross estate value), age of the decedent, and taxability of the estate; that is, whether or not an estate tax was due before the application of credits. Taxability is naturally a categorical variable. While age is discrete, it can take on over 100 values, thus age categories, similar to the categories used in constructing sampling strata, were used as dummy variables, as were size categories. The categories were chosen to reflect marginal changes in late-filing behavior based on exploratory analysis. Precise category boundaries were then adjusted due to the desire to have them, when possible, match sampling stratum boundaries, and the need to have sufficient numbers of late-filing events in each cross classification (taxablilty × age × size) to support modeling.

3.1 Survival Analysis

Survival analysis, or time-to-event modeling, is a well-known technique for measuring the probability that some event (death in its original application) will occur within a given time period. It has been widely used since to model more general time-to-event problems. The survivor function estimates the probability of an event occurring at or after some time *t*. In this context, the event of interest is the filing of an estate tax return, and "survival" equates to making it to the end of the sampling period cut-off (3 years) without filing an estate tax return.

One method for forming such a model is Proportional Hazards (Cox) regression. Cox regression is a widely accepted type of survival analysis model. It allows the use of covariates to help explain differences in times to some event for different observations. For the estate tax study, age of the decedent and size of the estate are both important predictors of time to filing. Cox regression can also handle other important features of the estate tax study data.

In order for an estate to come into existence, someone must die. Prior to his or her death, and the formation of the estate, there is no risk of an estate return's being filed. SOI conducts a study of estates of decedents who die in every third year. Since the dates of death are distributed throughout the reference year, estates are formed and become subject to filing at different times. This is similar to a study of, say, cancer treatments, where subjects may enter the study at time of diagnosis and thus many subjects may become part of the study cohort at different times. The phenomenon of some subjects' beginning to experience positive probability of an event's occurring at a later time than others is called "delayed entry," and the observations for those subjects are referred to as "left-truncated." Cox regression can handle left-truncated observations.

Using Cox regression, the authors estimated the parameters of the survivor function conditional on the values of the covariates. For every adjustment stratum (shown in Table 1), the authors fit a model to the estate tax study year-of-death 2001 population data. In order to do this, the authors analyzed all of the possible combinations of the selected covariates for each stratum, keeping the best set of significant covariates for each stratum. The authors also used the year-of-death 2004 sample file to create a vector of all

three covariates for each return. The authors then used the covariate vectors from the 2004 sample to predict a set of survival probabilities.

Table 1: Definition of Categories of Total Gross Estate and Age

Variable Name	*Lower Bound*		*Covariate*		*Upper Bound*
ageCats0	0	≤	Age	<	40
ageCats1	40	≤	Age	<	65
ageCats2	65	≤	Age	<	70
ageCats3	70	≤	Age	<	75
ageCats4	75		Age		or older
sizeCats0	$1.5 million	≤	Total Gross Estate	<	$2.0 million
sizeCats1	$2.0 million	≤	Total Gross Estate	<	$3.0 million
sizeCats2	$3.0 million	≤	Total Gross Estate	<	$5.0 million
sizeCats3	$5.0 million	≤	Total Gross Estate	<	$10.0 million
sizeCats4	$10.0 million		Total Gross Estate		or more

3.1.1 Survival Analysis Results

Table 2 presents new population estimates derived using the survival analysis approach as well as comparisons to known population totals and estimates using previous adjustment methods. The survival analysis model overestimated number of returns by about 6.5 percent and total gross estate by 10 percent.

Table 2: Year-of-Death 2004 Population Totals and Sample Estimates with Adjustment Factors Modeled Using Survival Analysis

Weight Adjustment Method	*Number of Returns*	*Percentage Difference*[1]	*Total Gross Estate ($ Millions)*	*Percentage Difference*[1]
Population total	**41,922**	**n.a.**	**149,430**	**n.a.**
Unadjusted estimate	40,453	-3.50	147,163	-1.52
Woodburn (1992)	40,785	-2.71	148,199	-0.82
Raub (2007)	40,867	-2.52	148,502	-0.62
Belmonte *et al.* (2010)	44,680	6.58	163,942	9.71

[1]Percent difference from known population total

The overestimation of both number of returns and total gross estate indicate that non-proportional hazards were not ignorable. The models were fit with time-dependent covariates to adjust for the effect of time on the effects of the different covariates. Many of the time-dependent covariates were highly significant. Also, their associated hazard ratios were greater than one, indicating that hazard, or risk, of filing increased as time passed. By ignoring the violation of proportional hazards, the hazards across time were essentially "averaged over". This led to an underestimation of hazard, resulting in survival probabilities for late-filed returns higher than acceptable for the desired outcome.

3.2 Logistic Regression

Filing before or after the designated sampling cutoff can be modeled as a binary response variable. Logistic regression is a commonly used method for predicting the proportion of times an event occurs in a number of trials conditional on the values of some explanatory covariates [2, 3]. As in the previous model, the selected covariates were size of the estate (again, measured by the total gross estate value), age of the decedent, and taxability of the estate. Definitions of the selected categories are shown in Table 3.

Table 3: Definition of Categories of Total Gross Estate and Age

Variable Name	*Lower Bound*		*Covariate*		*Upper Bound*
ageCats0	0	≤	Age	<	40
ageCats1	40	≤	Age	<	65
ageCats2	65	≤	Age	<	70
ageCats3	70	≤	Age	<	75
ageCats4	75		Age		or older
sizeCats0	$2.0 million	≤	Total Gross Estate	<	$3.0 million
sizeCats1	$3.0 million	≤	Total Gross Estate	<	$5.0 million
sizeCats2	$5.0 million	≤	Total Gross Estate	<	$10.0 million
sizeCats3	$10.0 million		Total Gross Estate		or more

3.2.1 Logistic Regression Results

Table 4 shows the analysis of maximum likelihood estimates. All categories of all covariates are highly significant. Model development was guided in part by residual analysis, influence measures, and goodness-of-fit tests, but, as this paper is primarily concerned with good predictions and not explanation, these are omitted here.

Table 4: Analysis of Maximum Likelihood Estimates

Parameter[1]	*DF*	*Estimate*	*Standard Error*	*Wald Chi-Square*	*Pr > Chi-Sq*
Intercept	1	-2.4574	0.1316	348.8	< 0001
ageCats1	1	-0.5127	0.1311	15.3	< .0001
ageCats2	1	-0.7958	0.1385	33.0	< .0001
ageCats3	1	-0.9031	0.1356	44.3	< .0001
ageCats4	1	-1.3849	0.1286	116.0	< .0001
sizeCats1	1	-0.1191	0.0400	8.9	0.0029
sizeCats2	1	-0.2640	0.0525	25.3	< .0001
sizeCats3	1	-0.5813	0.0786	54.7	< .0001
Taxable	1	-0.1385	0.0413	11.2	0.0008

[1]The effect of the first category of each of the dummy variables for Age and Total Gross Estate is reflected in the Intercept.

Results from this method were quite good. Table 5 presents new population estimates derived using the logistic regression model as well as comparisons to known population totals and estimates using previous adjustment methods. This method produced an excellent estimate of total number of returns, the predicted value for which the method

was designed. Additionally, the method resulted in a reasonable estimate of total gross estate.

Table 5: Year-of-Death 2004 Population Totals and Sample Estimates with Adjustment Factors Modeled Using Logistic Regression (for Returns with Total Gross Estate of $2.0 million and above)

Weight Adjustment Method	*Number of Returns*	*Percentage Difference*[1]	*Total Gross Estate ($ Millions)*	*Percentage Difference*[1]
Population total	**28,355**	**n.a.**	**161,007**	**n.a.**
Unadjusted estimate	27,701	-2.31	159,330	-1.04
Woodburn (1992)	27,926	-1.51	160,245	-0.47
Raub (2007)	27,981	-1.32	160,582	-0.26
Belmonte *et al.* (2010)	28,315	-0.14	162,213	0.75

[1]Percent difference from known population total

4. Future Steps

Estimates for the Estate Tax study benefit from a small adjustment to account for late-filed returns. As the research shows, logistic regression can be a useful method for calculating such adjustment factors. Results from logistic regression models are encouraging for the future development, assessment, and potential application of such models to adjust population estimates for other SOI studies. The authors recommend that efforts to develop similar models for each of the Private Foundation and Exempt Organization studies be undertaken as soon as possible.

5. Acknowledgements

The authors would like to acknowledge the management of the Statistics of Income Division, Internal Revenue Service, for their support and encouragement of this work. The views expressed in this paper are those of the authors and should not be attributed to the Internal Revenue Service or the U.S. Department of the Treasury.

6. References

[1] Raub, B., C. Belmonte, P. Arnsberger, M. Ludlum. The Effect of Late-Filed Returns on Population Estimates: A Comparative Analysis. In *JSM Proceedings*, Section on Survey Research Methods. Alexandria, VA: American Statistical Association, 2009.

[2] Agresti, A. *An Introduction to Categorical Data Analysis*, John Wiley & Sons, New York, NY, 1996.

[3] Stokes, M., Davis, C., Koch, G. *Categorical Data Analysis Using the SAS System*, SAS Institute, Cary, NC, 1996.

The Effect of Content Errors on Bias and Nonsampling Variance in Estimates Derived From Samples of Administrative Records

Barry W. Johnson and Darien B. Jacobson

Barry W. Johnson, Statistics of Income RAS:S:SS, P.O. Box 2608, Washington, DC 20013-2608

Key words: Bias, Non-sampling error

The Statistics of Income Division (SOI) of the Internal Revenue Service (IRS) uses a number of methods for ensuring the quality and integrity of the data it produces for tax administration research. As a first line of quality assurance, codes and mathematically related data items are extensively tested as SOI employees enter them into computer databases. In addition, for a sub-sample of returns selected and processed in most studies, SOI assigns a second employee to reenter and edit the data. Values from the first and second edit are then computer-matched. A supervisor resolves discrepancies discovered during the match. The original value, second value, and correct values are all collected as a part of the quality review system, as are a set of codes that describe the cause of the error, in broad categories.

This paper will use quality review data from Federal estate tax returns (Form 706) selected into the Calendar Year 2002 SOI Estate Tax Study to estimate the effects of non-sampling error on estimates derived from the final data file.

Background

The Federal estate tax is levied on estates for the right to transfer assets from a decedent's estate to its beneficiaries; it is not an inheritance tax. A Federal estate tax return must be filed for every U.S. decedent whose gross estate, valued on the date of death, combined with certain lifetime gifts made by the decedent, equals or exceeds the filing threshold applicable for the decedent's year of death. A decedent's estate must file a return within 9 months of a decedent's death, but a 6-month extension is usually granted.

All of a decedent's assets, as well as the decedent's share of jointly owned and community property assets, are included in the gross estate for tax purposes and reported on Form 706. Also reported are most life insurance proceeds, property over which the decedent possessed a general power of appointment, and certain transfers made during life.

Expenses and losses incurred in the administration of the estate, funeral costs, and the decedent's debts are allowed as deductions against the estate for the purpose of calculating the tax liability. A deduction is allowed for the full value of bequests to the surviving spouse. Bequests to qualified charities are also fully deductible.

Data Description

The 2002 SOI Estate Tax Study was a stratified, random sample of returns filed in Calendar Year 2002 and was the second year in a 3-year study of Federal estate tax returns filed 2001-2003. The sample was designed for use in both estimating tax revenues in all 3 calendar years and personal wealth holdings for 2001 decedents. The 3-year sample period was devised to ensure that nearly all returns filed for 2001 decedents would be subjected to sampling, since a return could be filed up to 15 months after the decedent's death. The design had three stratification variables: size of total gross estate plus the value of most taxable gifts made during the decedent's life, age at death, and year of death. The year-of-death variable was separated into two categories, 2001 year of death and non-2001 year of death, in order to facilitate studies of 2001 decedents. Returns were chosen before audit examination and selected using a stratified random probability sampling method. A portion of the sample was selected because the ending digits of the decedents' Social Security Numbers (SSN) corresponded with those in the 1-percent Social Security Administration Continuous Work History Sample. However, the majority of returns were selected on a flow basis using the Bernoulli sampling method.

The sampling mechanism was a permanent random number based on an encryption of the decedent's SSN. Sample rates were preset based on the desired sample size and an estimate of the population. Sampling rates ranged from 3 to 100 percent, with more than half of the strata selected with certainty.

Data collection for the 2002 Estate Tax Study was conducted at the IRS Cincinnati Submission Processing Center. Employees entered the data from the estate tax return into a database using a Graphical User Interface (GUI) data entry system. Nearly 100 distinct data items were captured, with some balance sheet assets recurring hundreds, even thousands, of times, as assets were allocated to 32 different categories, such as stocks, bonds, and real estate. Tax returns ranged in size from a dozen to many thousands of pages, including appraisals, investment account listings, and legal documents. Tests embedded in the data entry system were used to validate entries and to ensure that mathematical relationships among variables were correctly preserved. There were more than 200 validation tests performed on each tax return included in the 2002 study.

While embedded testing can assure that codes are correct within a given range of values and that fields are mathematically consistent, many of the decisions that employees make when transforming tax return information into statistically usable data are not easily tested. For example, while several codes may be valid, determining the best code to describe a particular taxpayer's behavior or characteristics cannot always be automated. To address this problem, SOI developed a double entry quality review system. This system is a valuable tool for measuring both individual employee performance and overall data quality.

Quality Review System

A subsample of returns in the 2002 Estate Tax Study was subjected to additional review for quality assurance purposes. Returns were included in the quality review (QR) subsample through two different mechanisms, 100-percent review and product review. The 100-percent review consisted of all returns that were edited while an employee was in training. Product review was selected after the training period had been completed, and it comprised a 10-percent random sample of each employee's work. The product review sample was selected on a flow basis method using a pseudorandom number called the Transform Taxpayer Identification Number, or TTIN. The TTIN is a unique random number that is generated by mathematically transforming selected digits of the decedent's Social Security Number. The TTIN was then compared to the sample number, which represented the sample rate, in this case 10 percent. If the TTIN was less than the sample number, then the return was selected for product review.

Under the double-entry quality review system, one return was entered into the computer system twice by two different employees. The first employee did not know that a return was selected for review until after the first edit was complete, and the second employee was not allowed to see the first employee's entries. Therefore, each return had two versions in the database, the first edit and the second edit, and each was entered independently of the other.

When both employees finished editing a return, the computer compared the values from the original and QR versions. In some cases, the two versions matched perfectly; so, the return was released from the system, and the first edit data was treated as final and stored for later analysis. However, if mismatches between the two versions occurred, the discrepancies were stored in a separate data table to be reviewed by a supervisor.

The supervisor reviewed the discrepancies and charged the errors, assigning two codes to each discrepancy--one to identify the incorrect value and the other to describe the cause of the error. A discrepancy code was assigned to the error to explain which version was considered incorrect. Discrepancy codes were assigned to one of the following: the first version, the second version, both versions, or neither version. An error was assigned to both versions if both of the employees entered or interpreted the information from the return incorrectly. In this case, the supervisor was also required to supply the correct data value. In some cases an error was not assigned to either version, usually when the discrepancy was the result of a data processing peculiarity and not a true database error. After the error was assigned a discrepancy code, a numeric error resolution code was assigned to describe why the entry was incorrect. Error resolution codes indicate situations such as spelling errors, incorrect money amounts, or incorrectly assigned codes.

Once the supervisor reviewed all the discrepancies, each employee was given a list of the discrepancies, along with the discrepancy and error resolution codes, so that any first edit errors detected during quality review could be corrected prior to considering return processing complete. The feedback from the review also enabled employees to learn from their mistakes on each return and carry this knowledge into the editing of other returns. In the end, there is a database consisting of a table that includes all the values from the second edit of the return as entered, a quality review table containing a record of each discrepancy between the first and second edits (along with codes indicating who made the error and why), and a final data table containing the correct version of the return data that will ultimately be sent to customers.

For this paper, only a portion of the quality review data was used for analysis. First, data that were collected during periods of training, 100 percent review, were excluded. Second, only errors that were charged to the first edit or to both edits, meaning that the error required a correction to the final data set, were retained. This was done because these errors are more representative of errors that remain in the roughly 90 percent of the 2002 estate tax sample that was not selected for quality review. Third, errors that reflected idiosyncrasies related to the edit process itself, and not true data errors, were eliminated.

Empirical Results

Quarterly accuracy rates for each employee who worked on the Estate Tax Study for 2002 were generated using the product review data (see Figure 1). These rates were calculated using the number of returns that had at least one error charged to the first edit divided by the total number of returns that had been selected for quality review. The accuracy rates for all of the employees are not very high. However, these rates are a return level measure; any return with one or more errors is considered incorrect. The Form 706 includes an average of 150 data entry fields, while

complex returns can have more than a thousand entries; so, the probability of making just one mistake is very high. In fact, the average number of errors for each return is only 6.3.

Traditionally, supervisors have focused quality improvement efforts on those fields that are in error most frequently. By looking at the occurrence of variables *ex-ante*, using the first edit data, and *ex-post*, using the final corrected data file, it is possible to identify the frequency of original edit errors in the quality review sample. Figure 2 shows the percent changes in frequencies for variables on the file; each diamond represents a different variable. Frequencies change because many variables on the file represent balance sheet items, assets like stocks, bonds, mutual funds, and various types of real estate, which are not necessarily present in each decedent's portfolio. When an asset is incorrectly classified, not only does it change the dollar value of estimate, it also changes the frequency of occurrence of that particular attribute or asset type in the population estimates. This can be particularly problematic if the asset is of special interest to researchers. For example, there has been much discussion in the press about providing estate tax relief to small business owners. Errors that either under- or overcount the number of estates that have small businesses could have an impact on this debate. The percentages shown on the graph represent the aggregate correct frequency in the overall quality review sample, less the aggregate number originally reported, divided by the correct number. Negative percentages indicate cases where an asset was incorrectly included on the first edit. For example, the first employee may have incorrectly classified a balance sheet entry as a publicly traded stock, while the second employee may have

Figure 1: Employee Accuracy Rates

	Accuracy Rates			
Employee	Quarter 1	Quarter 2	Quarter 3	Quarter 4
17000	46.3%	23.9%	41.7%	21.7%
17100	25.0%	0.0%	0.0%	0.0%
17200	29.2%	30.8%	31.9%	40.0%
17300	57.1%	100.0%	91.7%	33.3%
17400	52.1%	28.6%	50.0%	37.9%
17500	44.4%	24.1%	54.8%	0.0%
17600	42.2%	51.9%	33.9%	46.2%
17700	41.9%	28.6%	39.3%	34.5%
17800	49.1%	25.0%	58.5%	45.6%
17900	52.3%	34.3%	59.0%	50.0%
17001	23.1%	34.2%	18.6%	44.7%
17002	39.2%	33.3%	36.2%	45.0%
17003	22.9%	20.7%	37.8%	29.1%
17004	34.2%	31.6%	22.0%	72.7%
17005	30.8%	0.0%	0.0%	37.9%
17006	26.5%	27.7%	41.4%	42.9%

correctly classified it as a mutual fund invested in a mix of financial assets. The percent changes in frequencies are generally close to zero, but there are some notable outliers.

Figure 2: Percent Change in Frequencies, Original and Final Edits

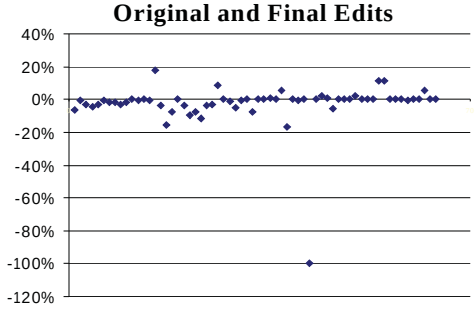
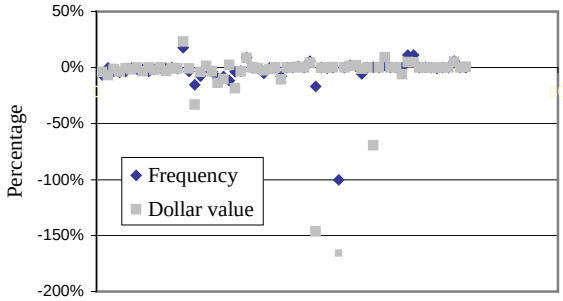


Figure 3 shows percentage changes in dollar amounts between first and second edits overlaid on the frequency differences shown in Figure 2. Each point represents a single variable on the file. While the pattern for the dollar differences is similar to that of the frequencies, with many differences close to zero, the magnitude of the dollar differences is larger for several variables. There are two variables for which the original entries resulted in aggregate dollar values that were overstated by roughly 150 percent. This highlights the potentially large effects on final estimates that can arise from even one large dollar value error, especially for variables that are not widely distributed in the overall population. Thus, it is important to monitor both the size and frequency of data entry errors.

Figure 3: Percent Change in Dollar and Frequency Values, Original and Final Edits



Unweighted error statistics are clearly useful for monitoring data quality and assessing opportunities for operational improvements during a study period. However, since the SOI study of Federal estate tax returns is based on a stratified random sample of the filing population, the effect of data entry error on final population estimates derived from this sample will vary inversely with the selection rate associated with each return. Using appropriate sample weights, it is possible to use the 10-percent QR sample to estimate the effects of data entry errors on population estimates derived from the remaining 90 percent of the returns in the final

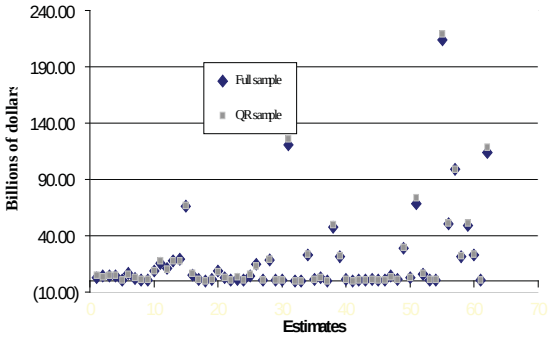
SOI data file that were not subjected to double-entry quality review. Weighted estimates provide a different perspective on the effects of nonsampling error due to the nature of the underlying estate study sample and the fact that the financial characteristics of estate tax decedents vary greatly among age and wealth classes. For example, younger decedents and those with large estates are selected into the estate tax sample with certainty and comprise more than 40 percent of the total sample file. Both groups of decedents are more likely to have had portfolios that are more complex and, thus, more subject to data entry errors than their either less wealthy, or older, cohorts. This is because many older wealth holders convert their portfolios to assets that produce tax-preferred income, usually resulting in returns that contain fewer business arrangements, which are more difficult to classify than market assets. Because the quality review sample is not stratified, weighted estimates will provide a more balanced measure of the overall effects of data entry errors on final estimates. Weighted estimates for the quality review sample were generated by using the design-based weight from the stratified estate study sample (W_s), multiplied by a quality review weight (W_q). The quality review weight itself was developed by first post-stratifying the quality review samples within the original selection strata as indicated below[1]:

Final Weight = $W_s * W_q$
Where $W_s = N_i/n_i$
Post-Stratification: $W_q = n_{if}/n_{qif}$

For some strata, the quality review sample was either zero or too small to create a post-strata cell. For these cases, strata were collapsed across age categories so that estate size classes were preserved.

Figure 4 shows full population dollar value estimates from the quality review data using the post-stratified quality review weight and compares them to population estimates using the full weighted estate study sample. Each pair of data points represents a different variable on the file. The quality review data estimates for each variable are denoted by the gray squares, and the full sample estimates are denoted by the black diamonds. For most variables, the QR sample estimates are larger than the population estimates from the full estate sample, indicating that the QR sample introduces a positive bias. This bias arises because the QR sample is a simple random sample of a stratified sample that favors large dollar value returns. In such cases, ratio raking can often be employed to decrease the bias; however, in this case, the QR sample size was insufficient in the lower gross estate size classes.

Figure 4: Full Sample vs. QR Sample Estimates



While the weighted QR data estimates are somewhat biased due to the design of the sample, they still provide an important indication of the effects of data entry errors on final estate tax sample estimates. Figure 5 shows weighted and unweighted estimates of aggregate differences between original and final values of both frequency and dollar value estimates for selected variables. A negative value means that a variable was over represented in the original, uncorrected data, and a positive value means it was originally underrepresented. Weighted results rank errors differently for some of the variables. For example, errors in classifying noncorporate business assets had a much greater impact on final weighted estimates than would have been evident had the analysis been limited to examining the unweighted QR data. Conversely, the unweighted QR data implied that the effects of errors on estimates of farm real estate were greater than they are in the final, weighted estimates. Clearly, using weighted estimates, along with the unweighted quality review data, provides a more

Figure 5: Differences between First and Final Edits

Data Element	Frequency	Dollar Value
Noncorporate Businesses	**-11.00%** *-5.29%*	**-5.79%** *-3.55%*
Closely held stock	**-3.06%** *-3.42%*	**-1.01%** *-0.71%*
Real estate	**6.70%** *6.82%*	**7.34%** *6.17%*
Farm land	**-0.91%** *-1.95%*	**-1.09%** *-3.66%*
Funeral expenses	**0.25%** *0.09%*	**0.15%** *0.04%*

Values in *italics* are unweighted estimates

[1] The subscript "if" signifies that certain reject returns were removed from the estate study sample prior to post-stratifying.

balanced method of assessing where to focus data quality improvement efforts.

Figure 6 compares the weighted percent differences between original edit estimates and final, corrected estimates with coefficients of variation (C.V.) from the full estate tax study sample in order to relate the sampling and nonsampling variances associated with selected fields. For some estimates, such as the values for noncorporate businesses and publicly traded corporations, the nonsampling error attributable to data entry is much greater than the sampling variance. For others, such as estimates of stock in closely held or untraded corporations and farm land, the sampling error, represented by the C.V., is actually greater than the nonsampling error attributable to data entry errors, indicating that data entry errors are not a significant cause of additional variance in the estimates. Fields for which nonsampling error is relatively large provide opportunities for future data quality improvement efforts.

Figure 6: Data Entry Error vs. Sample Variance

	Frequency		Money Amount	
Data Element	**% diff**	**C.V.**	**% diff**	**C.V.**
Non-corporate businesses	-11.00%	4.45%	-5.79%	3.89%
Publicly traded stock	15.02%	.78%	20.00%	1.17%
Closely held stock	-3.06%	3.47%	-1.01%	2.18%
Real estate	6.70%	1.92%	7.34%	2.19%
Farm land	-.91%	4.34%	-1.09%	4.68%
Funeral expenses	.25%	.57%	.15%	1.19%
Spousal trusts	4.25%	2.97%	1.29%	1.58%

Conclusion

There is much to be learned through careful analysis of the data generated by SOI's double-entry quality review systems. The results of these analyses can be used to improve data collection systems and enhance worker training. Information on nonsampling error should also be useful to data users who could use data quality metrics to more accurately interpret economic modeling results and to ultimately build models that are more robust.

This analysis, however, revealed that the database format and the type of data that are collected from the quality review samples make certain types of analysis difficult, if not impossible. While a complete copy of the second edit is saved for all QR returns, the original, uncorrected first edit values are not saved when first edit errors require corrections. Information on discrepancies is kept in all cases, but, because corrections can involve changing any number of related fields, it is difficult to reconstruct exactly the first employee's original entries. If more sophisticated analysis is desired, including the study of secondary errors that arise as a result of a primary data entry error, archiving a complete copy of the first edit, along with associated error reason and discrepancy codes, should be considered.

It is also important that supervisors apply error reason and discrepancy codes consistently. All too often, discrepancies are resolved by several different supervisors. Some, especially those serving in a temporary capacity, may feel a great deal of peer pressure to avoid assigning errors to individual employees, even in cases where the assignment of an error would not directly impact employee performance appraisals, such as when an error is attributable to lack of clarity in editing instructions. This inconsistency makes it difficult to measure the extent to which errors exist and to learn of ways to avoid them in the future.

Related to this problem is that the measure of employee performance currently in place is not adequate. It is simply unfair to use a return level measure of accuracy when the difficulty of the work is so variable across returns. A more balanced measure would relate the number of individual errors an employee makes to the number of fields he or she actually edited, thus giving full consideration to the number of edit decisions that were made on each return.

Finally, there are sample design issues that became apparent from this analysis. The QR sample is biased and could be improved by taking into consideration the underlying structure of the estate tax study sample design. Even this would not provide coverage of variables that are relatively rare, but perhaps important, in policy debates. To address this problem, samples could either be increased or targeted to include more returns with important characteristics, such as those filed for small business owners, or returns that, because of the types of entries made during first edit, are more likely to contain significant problems. Samples could also vary with worker skill levels. One possibility would be to develop a system that sets a weekly QR sample rate for each individual employee based on individual rolling average accuracy rates. Sample rates could be set automatically based on preset performance standards. Automating the process would avoid putting supervisors in the awkward position of having to 'punish' poor performers with additional oversight, making it easier to match feedback and training efforts to performance levels.