SATISFYING THE NEED OF INCOME POLICY MODELERS WHILE PRESERVING THE RELIABILITY OF DESCRIPTIVE STATISTICS

Susan Hostetter and Karen O'Conor, Internal Revenue Service

During the last few years, most Federal agencies have initiated improvements covering all areas of statistical methodology. The recent emphasis on quality methods, led by Deming (1986) and Juran (1988), is evident in the nature of changes in production of statistics throughout the Federal statistical community. Along the same lines as the recent cognitive laboratories, the statistical community has developed new cognitive-type methods of thinking or approaching new, and sometimes old, problems. (Dippo and Herrmann, 1991).

One key area of concern is how to meet the needs of the data user. Frequently, when data users become involved in specifying what they want, the bulk of the data are already processed and available; so, the user, in effect, walks down the shopping aisle, selecting which pre-established products he or she can best apply. Unfortunately, this approach rarely provides the best data to meet the analysts' needs. It was that realization that led the Statistics of Income Division, Internal Revenue Service (IRS) to work jointly with its major customer, Treasury's Office of Tax Analysis, to design a new sample for their tax policy microsimulation model.

This paper will describe that effort, with a focus on the process IRS used to improve our data products and make them more "fit for use." We will describe the problems that Treasury and Congress's Joint Committee on Taxation (Joint Tax) identified by using our data and provide some information about what prompted Treasury to want the sample changed. Treasury's request addressed the very core of the Individual tax statistics program -- the selection of returns. To set the stage, we'll give some background on how IRS has done its work in the past and about how that is changing since this effort began. The paper will review the development and results from the original Planning Team and will describe the implementation of some of the plans it recommeded. Finally, we'll bring you up to date on current and future work.

BACKGROUND

IRS' Statistical Culture

Although IRS is not a major data collector in the large-survey sense, it is the largest producer of data from administrative records. However, working in an administrative environment places important constraints on both the data producer and the data user. For instance, because of the IRS statistical culture, many of the design variables, such as the (1) **universe definition**, (2) **frame definition**, (3) **unit definition**, (4) **stratifying variables**, (5) **sample rates**, and (6) **collection mode (such as personal visit, mail, or telephone)** are already set or determined before IRS staff begin designing the sample. When we first started the redesign, only stratifying variables and sample rates seemed available for change.

Our universe is the current year's tax filers; our sample frame is the Master File of these tax filers; the sample unit is the tax return; and our collection mode is represented by identifying returns and stripping off tax information from those returns. Hence, our options are unusually rigid -- we have a lot of information from the tax returns, but we cannot clarify or improve our information, regardless of the problem. (For more information on quality efforts in the administrative and survey environments, see O'Conor et al., 1990.)

Finally, unlike some of the larger statistical agencies -- e.g., Census and the Bureau of Labor Statistics -- who have a diversity of major customers, IRS has a narrow customer base. Until we have met the needs of Treasury and Joint Tax, who support our budget, IRS cannot address the needs of other users. This is true even for major needs, such as those of the Bureau of Economic Analysis or our own IRS published tables for general public use.

All of these factors presented special problems when we began to redesign the sample of individual income tax returns.

Tax Policy and Income

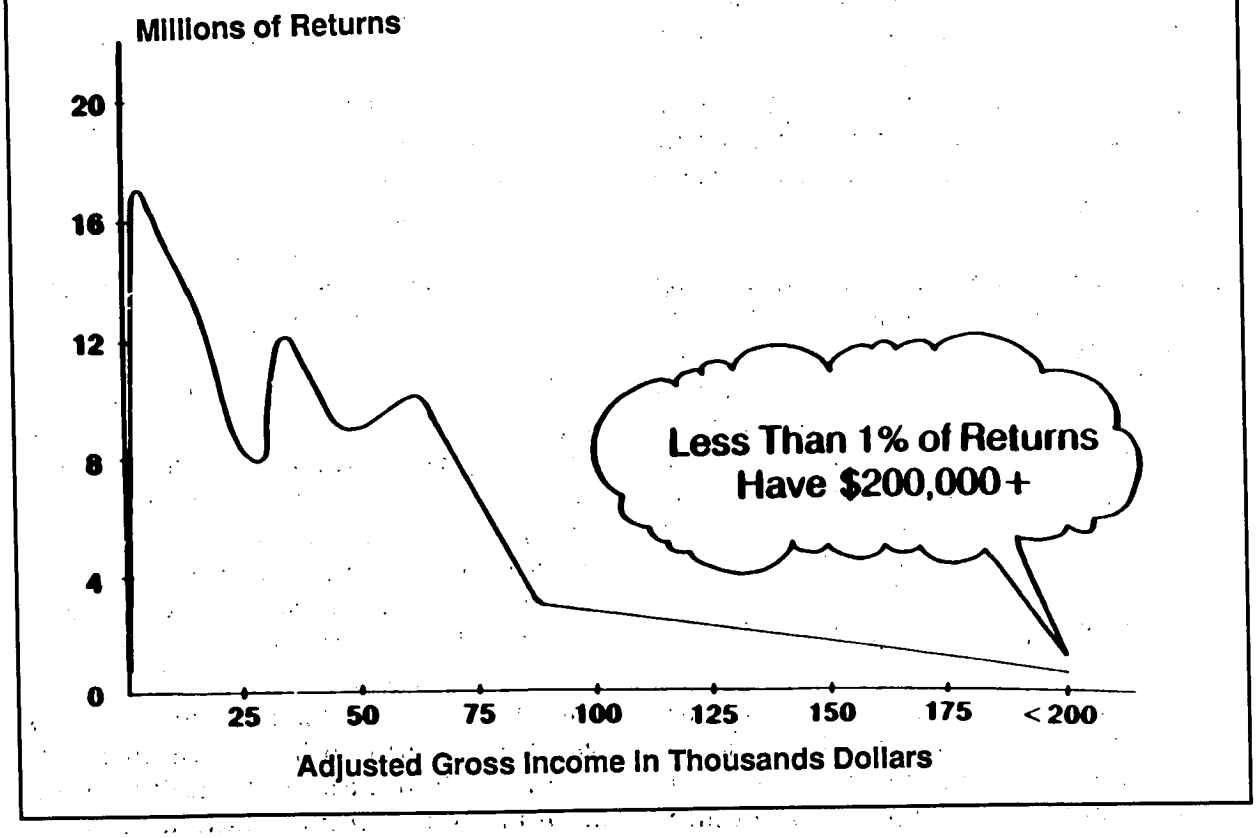
Why was the redesign necessary? Well, tax policy is really about income -- how much and which kind gets taxed. Of course, income is not distributed across the population evenly at all. (Nelson, 1986 and Hostetter, 1987.) One of the particularly difficult tasks in our work is designing within the realities of this income distribution. As you can see in Figure 1, 99 percent of the individual returns are from $0 to $200,000. However, that remaining one percent of returns covers 13 percent of the income. Even more startling, the millionaires represent .06 percent of the returns, but have five percent of the income. These figures are based on Adjusted Gross Income (AGI), which is actually not representative of gross or spendable income. If the percentages were based on gross income or the selection amount for our new sample design, which is closer to gross income, the distribution would be far more skewed than that show in Figure 1. It is the individuals at the upper tail who have the most income protected from taxation and have the least percentage included in AGI. Naturally, these are the data of greatest interest to tax policy analysts.

Finally, we should mention two additional points that are important considerations in designing a sample of tax returns selected primarily on income. The IRS sample is heavily skewed to select many more high income returns than middle or lower income returns. We select the very high income returns at a 50 or 100 percent rate. Secondly, we see more changes (from year-to-year) downward between strata than upward, because we have many fewer of the lower income returns that might move up. (Schirm and Czajka, 1990.)

Motivation for Beginning a Sample Redesign

User Needs.--The agony that our two major users -- Treasury and Joint Tax -- suffered to prepare revenue estimates during the development of the Tax Reform Act of 1986 was like the starter on a car for the beginning of our process. Much of the specific difficulty in developing revenue estimates for the Tax Reform Act involved how you treat income as it relates to each tax policy proposal. Treasury and Joint Tax did not totally agree on using income sources other than those from tax returns, so their estimates of potential revenue for proposals were not always the same.

Figure 1.--Distribution of Returns by Adjusted Gross Income, 1989



Treasury initiated our research process to address their need to have a more reliable and comprehensive database -- "fit for their use." Unfortunately, unlike testing the effects of a medication on an illness, one of the problems in researching social policy issues is that we cannot conduct experiments -- all we can do is develop longitudinal observational studies. Luckily, in our case, that was a strength. In IRS, our area of control is such that our sample frame has extensive information that allows us to select units with enough precision that we can conduct narrowly defined observational studies. (Our Longitudinal Capital Gains Study that covers information back to 1979 is a good example.)

Technological Development.--In recent years no one can keep score or stay up-to-date with the newest computer innovation. The capabilities grow faster than we can accurately define ways to use them and, particularly, to document them properly. This "revolution," as they call it, has enabled a broad range of users access to microdata to develop their own estimates. By necessity, it has also made users more aware of the nuances of data relationships for specific inferences. They, in turn, are beginning to have opinions about the selection and processing of data. Since our two major users are provided the full microdata file to develop and use for tax modelling and revenue estimates, they have had ample exposure, for many years, to developing their own uses from microdata. Needless to say, they also had many ideas about how to improve the data and clearly wanted to play an active role in doing so.

Quality Environment.--One more motivating factor should be mentioned -- quality. In 1988, IRS, as an agency, adopted a Total Quality Management (TQM) philosophy, to the extent that all managers received TQM training and many also attended training in the Juran Quality Process, as well. This new way of thinking strongly influenced the way we addressed our user's concerns.

TREASURY'S TAX POLICY MICROSIMULATION FILE

Treasury and Joint Tax develop tax policy with the aid of microsimulation that combines the data available from tax returns with survey data from other statistical agencies. From this process they obtain what they call the Tax Model. (An overview of the Treasury Tax Model is given in Cilke and Wyscarver, 1987, and detailed documentation is provided in Cilke and Wyscarver, 1990.) The development of Treasury's Tax Model, shown in Figure 2, had the following three phases:

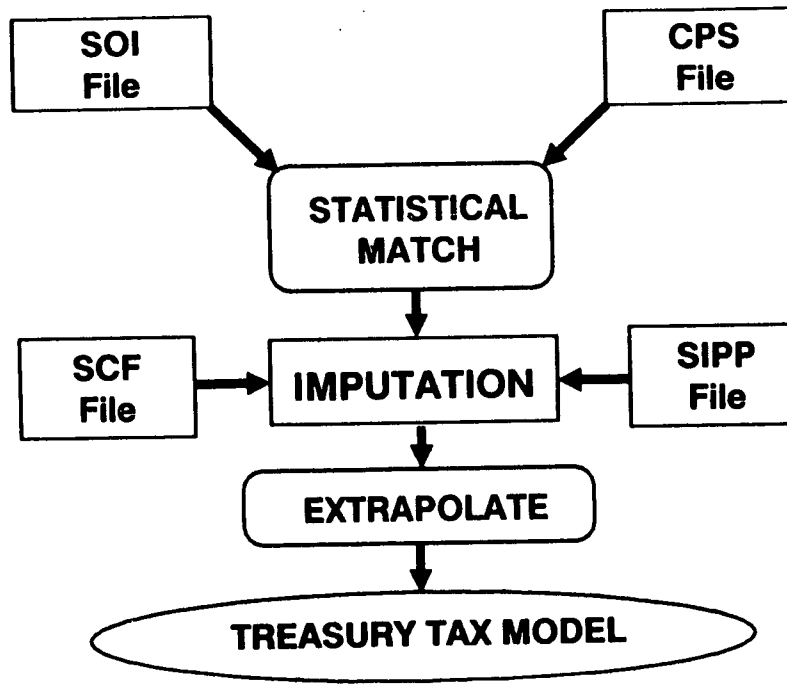
- First, there is a matching phase, using IRS's Statistics of Income (SOI) Individual Program microdata file as the base, or host, file and matching Current Population Survey (CPS) household individuals to the IRS taxpayers. This is matching, or linking, of statistically "similar" individuals, rather than exact record linkage of identical individuals. Treasury staff use age and income as determining variables in the matching process, and the match is constrained so the marginal distributions of the two populations are maintained. (Barr and Turner, 1978.)

- Second, Treasury imputes data from the Federal Reserve Board's Survey of Consumer Finances (SCF) and the Census Bureau's Survey of Income and Program Participation (SIPP) files.

- Third, after matching and imputation, Treasury uses the extrapolation process to forecast five years ahead. Since the data are, by this time, two years old, adjustments using more currently available data are also employed. (Remember, tax data are, by definition, one year old -- or covering the previous year -- when returns are filed and, hence, are first available to be retrieved.)

To most accurately complete these matching and imputation tasks Treasury's need to improve their file to tax family units instead of tax return units was based on two issues. First, the family economic unit more closely resembles the individual's level of well being and, therefore, provides a fairer basis for reviewing tax policy. Second, the match of IRS data to other sources to obtain income not reported on tax returns (the sources identified in Figure 2) is based on household units, and the tax family more closely resembles the household than does the return. In other words, all three of the surveys that Treasury uses to add data to the host SOI file by matching or imputation

Figure 2.--Department of the Treasury's Tax Model



are based on household units. While many statistical agencies may change their sample unit easily, or have different sample units for different surveys, in IRS, changing from a return-based to a tax family-based survey was a very major conceptual change.

Since tax returns provide most, but not all income, these additional sources and their treatment in processing is important to the success and accuracy of the Treasury Tax Model. Comparability of income among the surveys and the reliability of income measures are important in choosing which surveys to use in building the microsimulation file. Because confidentiality constraints prohibit exact linkage and the IRS file contains limited matching-type information for a statistical match, income is used as the key variable for matching or imputation among the SOI file and other surveys. Fortunately, there is a lot of commonality in the income definitions for all three of these surveys. Most of the new work by Arthur Kennickell (1991) and Janice Lent (1991) has led to improvements in comparability of income. The IRS effort to develop a panel of tax families will make the IRS matching unit more similar to the household unit used in the other surveys, and improving the treatment of income reported on tax returns will improve the quality of the primary matching variable. Thus, these improvements to data quality will also increase the usefulness of the policy microsimulation file.

EARLY PLANNING STAGES FOR NEW SAMPLE DESIGN

What was wrong with the IRS sample that had served us so well for 10 years? The first two changes requested were to develop the tax family and to identify a longitudinal sample or panel. These, of course, were more easily identified than implemented.

One of the initial problems with the old sample design was that Treasury staff felt that returns with

large amounts of negative income were not included in our sample. (Negative income refers to losses reported on returns, which -- for tax purposes -- are subtracted from reported positive income. A negative return is a tax return that has a larger total of negative income than of positive.)

They also said that including returns with positive income and those with negative income in the same strata -- as the old design did -- created strata with too much heterogeneity. Lastly, Treasury staff said we simply did not have enough of several kinds of returns that they needed to examine for tax policy implications.

The first step IRS took to address these issues was to establish a Quality Planning Team. Begun in 1987, it was the first Planning Team in IRS, and it met regularly for over a year. This initial effort was aimed at creating dynamic change in a technically complex program that was managed in a very large bureaucracy. Success overall may very well be due to the strengths of the original Planning Team. They were:

- the priority-giving support that Treasury provided;
- use of the Juran Quality Process, which emphasizes the inclusion of information and evaluation before decision-making;
- use of minutes of meetings as a planning tool;
- the decision to hold meetings off-site; and
- the diversity and broad representation included on the Team.

The Team included IRS National Office program staff, IRS Martinsburg Computing Center staff (where our sample is selected), IRS service center processing staff (where our editing is done), IRS Detroit Computing Center staff (where our final processing is completed), a statistical contractor (Mathematica Policy Research, Inc.), a Juran-trained facilitator, and last (but definitely not least) Treasury staff.

The processing experts were important to help us determine and define necessary processing methods for new ideas. They were instrumental to the eventual success, not only for their knowledge, but also because answers and interaction could be defined immediately, which, in turn, allowed us to make assumptions and continue planning. The customer representation was essential. They not only played an active role on the Planning Team, but also provided associated research to test the design, financial resources during a period of critical budget constraints, and overall support and visibility. In summary, the broad coverage gave us ownership of product where all the crucial implementation responsibilities were.

The results of the Planning Team's work were the definition and preliminary processing design for both the tax family unit and the panel and the agreement to, and schedule for, developing and implementing a new cross-sectional sample design. The rest of this paper will focus on these efforts and describe how the process worked. (Early ideas for developing a new income stratifier were described in Czajka, 1988.)

TRANSITIONAL IMPLEMENTATION

During the initial planning stage, Treasury staff emphasized strongly that they needed to have the identification of the panel and the building of tax families begin as soon as possible. Since planning began in 1987, the Planning Team decided to use the Tax Year 1987 sample as the base year for the panel. These returns would be filed and selected during 1988. The concept of the tax family unit was developed and the extremely sophisticated and difficult processing necessary to support these families was specified. Now, during this early phase, these plans lacked perfection, but, in spite of this limitation, they were implemented within a few months of beginning the planning phase. IRS refers to this work as the transitional implementation.

The Implementation Process

Obviously, such a complex system to address multiple needs was not easy to develop. It clearly represented much interaction and consensus. Originally, the implementation work was not separated from program production, but IRS management recognized the project's vulnerability and initiated a new organizational unit to shepherd the implementation and reinitiate planning. The resulting intermediary (or use transitional) product met some of our objectives, but by far the greatest value realized was in the

lessons we learned from the process. Following is a "lessons learned" list of problems that IRS suffered. Many, no doubt, are also common to new or dynamic efforts in other Federal agencies.

- **Data Capture System** -- IRS uses the identification of the primary taxpayer on tax returns only in processing, and our need to have returns identified by either primary or secondary social security number really challenged the programmers at the Martinsburg Computing Center.

- **Transition from Planning to Implementation** -- There is some loss of visibility in implementation that may translate into loss of support. It is difficult to do developmental or "quality-related" work along with production work.

- **Diluted Attention from the Top** -- The length of the process, as in the case of this sample redesign, may eventually contribute to reduced attention and, therefore, support from the top management.

- **Implementation of an Innovative Plan in an Institutional Climate** -- In general, the larger the organization, the more widespread is the need for changes to accommodate the new "invention," and frequently the more resistent the institution will be to such change. IRS is a large organization.

- **Support in Resources, Time, and Money**

- **Unexpected Additional Complexities** -- The widely publicized IRS contracting problems certainly fit this description.

- **Mistakes for First-Time Users** -- IRS complexity, and that of many statistical agencies (especially those using administrative records), actually contributes to making mistakes in first-time projects.

- **Conflicting Interests of Users**

REDESIGNING THE CROSS-SECTIONAL SAMPLE

With those lessons to guide us, in the spring of 1989 IRS began to work with Treasury staff to redesign the annual cross-sectional sample. In the past, we might have taken a "traditional" approach, in which samplers and data users spend several meetings determining users' needs. These are frequently described in the form of estimates required, or tabulation specifications, with a fixed reliability for each estimate. In this very typical scenario the participants would agree on the rather simplistic needs assessment, and the sampler would go off independently to design a sample that satisfies the requirements. Everyone would have done his job. This process is concise and efficient, if tabulations are the desired product.

Instead, our commitment to the new process led us down a different path. Unfortunately, the methods we used in the earlier planning phase of this project did not meet our needs for planning the sample design. We decided, instead, to go to the customer. We realized we needed to provide **a service -- not a product**. We worked with the customer as partners in the process and, as a result, the customer contributed to the quality. To be effective, we had to learn to listen naively to our customer. We must note here that we did not start out this phase so perfectly. We both -- Treasury and IRS -- learned together, by trial and error.

The Decision Process

The developmental team used a number of sample design tools during the year and a half process that included regular meetings to review research results. Most of the research was done by our contractor, Mathematica Policy Research Inc., in reaction to issues and questions from the previous meetings. In other words, at a typical meeting, we would make a few decisions and agree on what we needed to look at to make some more decisions. The decisions were iterative and based on data that we all reviewed. All questions and unreadiness were respected, regardless of the number of individuals involved. Throughout the process we controlled our changes to ensure that the quality of data would not be appreciably less than that of the old design. After all, this old sample design had served IRS well for 10 years.

After several preliminary decisions, we established initial new strata and test selection rates, which we used to simulate the first version of a new sample and compare to the old one. We used the 1987 old sample for simulating results from a new design -- there was good reason to believe the results would be

reliable, mainly because we expected about a 70 percent overlap between the old and new. Estimates were calculated by assigning a new stratum definition to each return and calculating estimates using the new strata. (A detailed comparison of the old and new sample is provided in Schirm and Czajka, 1991.) Among our principal goals were strengthening the sample of income components and the ability to obtain better coverage for certain demographic groups which are the subject of tax policy. (Hostetter et al., 1990, and Czajka and Schirm, 1990.)

Development of Preliminary Selection Rates for Testing the New Sample

Figure 3 compares the reliability of estimates from the new sample to those of the old sample, by looking at the coefficients of variation (CV's). Comparisons are also provided to the constrained optimum -- an estimate of the best CV that could be obtained for that estimate, using Neyman allocation based on that variable.

The "Old" column shows the CV's that result from the old design, which is optimized on Adjusted Gross Income (AGI) -- as you can see, for that variable, both the old design and the optimal CV's are 0.015. While such a low CV may seem unusual, the fact that we are dealing with a very large sample, largely based on AGI, meant that this result was to be expected. On the other hand, the rates for the new design are not the optimal rates for estimates of common characteristics, such as AGI, wages, interest, tax liability, or even itemized deductions. They do, however, result in an improvement for more unusual items, such as tax exempt interest, social security income, gross capital gains, gross partnership income, or tax preference items. In fact, of the 35 tax items we reviewed, the new design had lower CV's for 27. This was important for Treasury's tax policy microsimulation purposes, since most of the fields used to calculate the new income stratifier would be these more unusual data items.

With these CV's to quantify what we were look-

Figure 3.—Comparison of Coefficients of Variation by Sample Design

Income or Tax Item	Constrained Optimum	Sample Design	
		Old	New
Common			
Adjusted gross income/deficit	0.15	0.15	0.19
Salaries and wages	0.21	0.23	0.28
Taxable interest	0.72	0.97	0.98
Net capital gain	0.90	3.77	4.30
Taxable pension/Annuity income	0.91	1.42	1.42
Total tax liability	0.23	0.23	0.27
Total itemized deductions	0.36	0.46	0.49
Rare			
Dividends	1.09	1.50	1.37
Taxable Social Security income	0.93	2.19	1.98
Tax exempt interest	1.30	5.16	2.11
Gross short-term capital gain	1.67	3.58	2.06
Gross long-term capital gain	0.69	1.12	0.88
Gross partnership/S corp. income	0.98	1.39	1.13
Tax preference items	2.08	4.41	3.45

ing at, the developmental team was able to make another round of decisions about the stratifying variables needed to select the new sample.

Treatment of Negative and Positive Income

As you recall, Treasury planned to use the data to estimate the potential effects of tax law changes. Because negative returns are essentially different from the great majority of positive returns, Treasury, therefore, wanted the income stratifier broken into positive and negative components. Additionally, they were concerned about the treatment of sources of income used for selection and wanted net fields (where losses have been subtracted from gains or where expenses were deducted from profits) used as seldom as available data would allow. To further reduce the "blurring" caused by merging positive and negative income, Treasury wanted **all** positive income totalled separately from **all** negative income, to obtain the income stratifiers. Finally, a return with a negative income selection amount would be sampled at the same or higher rate as a return with positive income of the same amount.

As you can see in Figure 4, for the old design, where net capital gains was the component in the stratifying income, the CV's were better for the old than the new design. However, in the new design, where gross gains or losses were used, we can see a considerable improvement in the CV's of the new design. This is reflective of the incremental changes, where we studied the effect of many items one by one and looked at how each was used in data analysis. It should also be noted that employing gross amounts has an added importance for income sources, such as capital gains, since the gross gains and losses are most frequently used in recommending tax policy.

During this review and analysis period we also examined a number of tables that arrayed units in the new sample strata compared to the old sample strata. Stimulated by coverage descriptions of the old sample, Treasury staff was able to describe additional changes they wanted. For instance, they added requirements for strata with a 100 percent selection rate for returns with positive or negative income of $5 million or more and a 50 percent strata requirement for returns from $2 to $5 million.

What is an Interesting Return?

Still, the sample was not quite right for policy analysis purposes. Initially, Treasury knew they wanted a better selection of returns for tax policy microsimulation, but they had difficulty expressing the specifics necessary to develop a statistical design. They could articulate that they wanted better cover-

Figure 4. — Comparison of Coefficients of Variation for Selected Income and Tax Items by Sample Design

Income and Tax Items	Sample Design	
	Old	New
Net capital gain	3.77	4.30
Net capital loss	.25	.29
Gross short-term capital gain	3.58	2.06
Gross long-term capital gain	1.12	.88
Gross short-term capital loss	4.45	1.98
Gross long-term capital loss	5.16	1.66
Net other gain (Form 4797)	4.59	3.76
Net other loss (Form 4797)	6.31	4.90
Net farm profit	4.64	3.73
Net farm loss	3.22	2.97
Employee benefit programs expense	7.55	4.79
Depreciation deduction	2.17	1.21
Gross rent/royalty income	2.03	0.95
Gross rent/royalty loss	1.39	0.96

age of the elderly, of capital gains income, of wealthy taxpayers, and a number of other specific items, including those in the list below:

- Partnership Income,
- Itemized Deductions,
- Rents Received,
- Social Security or Pension Income,
- Over 65 Years of Age,
- Child Care Credit,
- Head of Household Status, and
- Unemployment Compensation.

To exemplify their concern for interesting returns, Treasury staff said they wanted the 25-year old, average looking taxpayer, who will become a millionaire at 45. At this point in the process, we had satisfied many of the concerns and predicted that we had met many of the goals. However, we still had the bulk of taxpayers in the group from $0 to $250,000 positive income. While this group is homogeneous, there are some returns that Treasury wanted more of for policy analysis. We continued our research and reviewed reams of tabulations with hours of group review and discussion.

Finally a light bulb went off -- we decided to identify our interesting returns as those that were **not uninteresting**. Coverage in the new sample for these and more characteristics was estimated thoroughly, using volumes of tables that compared the new coverage to that in the old design, where Treasury knew from previous use whether the results were what they needed.

Substrata Improvement for Sample Coverage

Figure 5 illustrates the impact of the substrata on the sample selection rates. IRS staff expects about 23,000 returns from $0 to $30,000 in our sample from the population of about 69 million returns. Before stratifying for interesting returns, all 23,000 would be selected at a rate of .033 percent. However, we estimate that only 3,000 of the sample would meet our definition of "interesting." After defining and stratifying to select more interesting returns, we increased the number of interesting returns to 8,000. (See Hostetter et al., 1990, for definitions of the substrata, as well as definitions of all the income strata.) Upon completing and fine-tuning this final stratification, we again reviewed our coverage of all identified aspects of the new design, including assuring ourselves that we had not omitted returns with large values for key variables. Then, we finally began implementation of this new sample design.

COMBINED CROSS-SECTIONAL AND PANEL WEIGHTING

As mentioned earlier, a major purpose of the redesign is to permit the selection of both longitudinal and cross-sectional samples of tax families -- and, further, to combine the two samples, weighting them together for estimates. In so doing, dependent returns -- on which the filer is also claimed as a dependent for purposes of an exemption on another return -- and nondependent returns -- all other returns -- are treated quite differently than in the past. Previously, all returns, regardless of their dependent status, were selected with probability based on their income. Now, for the panel, only nondependent returns are selected with probability; dependent returns become part of the panel only if they are linked to a sampled nondependent return on which they are claimed. On the other hand, some dependent returns are still selected with probability based on their income in the cross-sectional sample. Usually these returns have relatively low income and are selected at low sampling rates.

The panel was selected from the 1987 cross-sectional sample, with some minor adjustments. Both the panel and the tax family unit were implemented immediately by identifying the nondependent panel returns, identifying the dependents claimed on them, and adding their SSN's to a "tickler file" that was used at the beginning of the Tax Year 1988 processing year. At the end of each year, additional dependents claimed on panel returns and all the dependents claimed on cross-sectional returns were identified and, in a separate computer processing effort, their returns were selected from the IRS Individual Master File. This process is updated and continued each year. So, by mid-1990 we will have a file with all the returns filed by panel members and their dependents. From the beginning, Treasury challenged the experts by requesting a combined, weighted file of cross-sectional and panel tax families. This work is currently underway in IRS, with the aid of Mathematica Policy Research, Inc. The first preliminary combined weighted file covering Tax Year 1988 was

Figure 5. — Improvement Introduced by Substratifying on Interesting Returns

Substrata	Sampling Rate	Number of Returns Selected
	Before Substrata	
Uninteresting	.033%	9,000
Moderately interesting	.033	11,000
Interesting	.033	3,000
	After Substrata	
Uninteresting	.020%	5,000
Moderatelyinteresting	.030	10,000
Interesting	.080	8,000

provided to Treasury and Joint Tax late in 1991. The rest of this section describes one of the more interesting results from this long developmental process -- our methods for weighting the combined cross-sectional and panel file.

The actual linking of dependent returns to the parent returns becomes more complex each time you account for a tax complexity, a taxpayer error, or unusual filing patterns. Our current control is a Panel Identification Number that identifies the 1987 "parent return" from which any return was identified and selected. We are in the process of lengthy and detailed review of thousands of panel units with three years of returns for all those who filed. (A full description and explanation of panel and cross-sectional selection and the combining of these returns can be found in Czajka and Walker, 1989. Also in Czajka and Schirm, 1991, the weighting of the combined cross-section and panel samples is more fully described.)

Figure 6 illustrates weights that will be used in the new sample management as they relate to dependent and nondependent returns in cross-sectional and panel files. It is evident that IRS will be greatly increasing the number of weight fields. So much so that there have been some remarks concerning a control file to manage weights. The increase in the number of weights is directly related to IRS redesigning and restructuring the file to make it "more fit for use" for more varied users.

Figure 6 describes the panel segments which will be weighted and the types of returns which will be included in each of the weighted estimates. A "Yes" indicates that returns selected according to the description in the left column for that sample segment will be weighted. For example, panel-only returns will be weighted for both the panel and combined sample segments. An "N/A" indicates that returns selected as part of the group described in the left columns for that sample segment are not applicable. For example, panel-only returns will not be included in the cross-section sample segment. "Sometimes" appears in two blocks because there are three types of dependent returns and two types of combined sample segments.

The combined segment is either a **return-based** cross-sectional survey, where the panel segment nondependent returns supplement the cross-sectional segment, or it is a **family-based** cross-sectional survey, where the panel families supplement the cross-sectional segment families. This is further complicated by the two types of dependents included in the cross-sectional sample. The first group is dependents included because their returns were selected in the probability sample, and the second group is dependents included because they are claimed as dependents on selected nondependent returns. The panel has no dependent returns selected with probability; only dependents included because they were

Figure 6. — Weighting for Nondependent and Dependent Returns by Sample Segment

Return Selected for:	Sample Segment		
	Cross-Section	Panel	Combined
	Nondependent Returns		
Panel only	N/A	Yes	Yes
Cross-sectional only	Yes	N/A	Yes
Both	Yes	Yes	Yes
	Dependent Returns		
Panel only	N/A	Yes	Sometimes
Cross-sectional only	Yes	N/A	Sometimes
Both	Yes	Yes	Yes

claimed as dependents on selected nondependent returns.

In the combined, return-based survey, the dependents claimed on nondependent selected returns cross-sectional segment and all the panel dependent returns are not included in the sample, but the probability-selected dependents in the cross-sectional survey are included. On the other hand, in the combined family-based survey, the dependents claimed on non-dependent selected returns in both the cross-sectional and panel segments are included in the sample, but the probability-selected dependents in the cross-sectional survey are not included. The following outline may help to clarify these descriptions:

A Combined Return-Based, Cross-Sectional Sample Includes:

A. From the cross-sectional sample:
   - Nondependent returns
   - Dependent returns selected with probability

B. From the Panel:
   - Nondependent returns

A Combined Family-Based, Cross Sectional Sample Includes:

A. From the cross-sectional sample:
   - Nondependent returns
   - Dependent returns selected because they were claimed on nondependent returns

B. From the Panel:
   - Nondependent returns
   - Dependent returns selected because they were claimed on nondependent returns

The combined file will have the following four weights for producing cross-sectional estimates:

- **A cross-sectional return-based weight.** -- This weight will be used to produce estimates for cross-sectional and overlapping panel returns. The only dependent returns included in these estimates will be those selected with probability in the cross-sectional sample. Estimates based on these weights will be used to produce preliminary tabulations and a microdata file for IRS' primary users. These are based on essentially the same weighting methods we have used in the past.

- **A combined return-based weight.** -- This weight will be used to produce estimates for all cross-sectional returns selected with probability (both nondependent and dependent) and for panel returns that are nondependent (or were originally selected as nondependent and became dependent returns after selection). The change to dependent status might occur if an older taxpayer became the dependent of adult children. This group of returns would be used to produce IRS annual tabulations and special tabulations that are prepared for the Bureau of Economic Analysis.

- **A cross-sectional family weight.** -- This weight will be used to produce cross-sectional estimates of tax family units from the cross-sectional and overlapping panel returns with their dependents linked to the correct nondependent return. Dependents selected with probability in the cross-sectional sample would be omitted from such estimates. It is likely that a file based on these weights would be one of three major files used for tax policy modelling and revenue estimates, with emphasis on more current cross-sectional data.

- **A combined family weight.** -- This weight will be used to produce cross-sectional estimates of tax family units for nondependent returns in the cross-sectional sample and the panel. Again, in estimates for tax family units, the dependents selected with probability in the cross-sectional sample would be omitted. In addition, panel individuals who were originally selected as dependents, but who now have established a separate tax family unit would have a different family identification number than the original parent of the panel unit and would, therefore, not be linked to a weighted nondependent return. They would also be omitted. The combined (cross-sectional and panel) tax family file developed from these weights would be the second of the three major files used each year by Treasury and Joint Tax to develop revenue estimates.

In addition to these cross-sectional estimates, IRS expects to develop weights for panel returns that will be used to produce longitudinal estimates. The panel, linked for about five years, is the third major file that Treasury and Joint Tax will use to develop revenue estimates. While not as current, this file will reflect the changes in the reporting behavior due to tax law revisions, economic effects, or taxpayer behavior at the individual level.

CURRENT STATUS AND FUTURE PLANS

Where do we stand now?

- The IRS tax family units and longitudinal data have been in place since 1988, the second year of the panel. We recently delivered to our primary customers the first weighted, family-linked, combined cross-sectional and panel data file.

- With three years (1987 through 1989) of data available on returns with dependents, we are currently undertaking a large manual review of all panel units (three years of returns for nondependents and their linked dependents) with potential error. We are developing methodology to code changes, so that what we learn about these returns can be applied -- to dependents of cross-sectional returns -- on an annual basis.

- In August 1991 the new stratification design went into production successfully at the Martinsburg Computing Center and, subsequently, in all 10 service centers.

- Treasury and Joint Tax staffs met with us to design subpanels to meet various tax policy processing needs. We also discussed a future capital gains subpanel.

- In 1993, IRS plans to provide published estimates based on the combined cross-sectional and panel file.

- Finally, in 1994, IRS plans to begin on-line longitudinal editing. With information for three years available during the on-line editing process for each individual, reviewers will make more knowledgeable corrections, thereby increasing overall editing accuracy.

In conclusion, this whole process has clearly been quite intensive and demanding -- of time, resources, ideas, and effort. Nonetheless, we have found it very rewarding. It has not, however, reached an end. If we have learned nothing else, it is that the process of serving our customers needs continues to change as their needs change.

ACKNOWLEDGMENTS

The authors thank Wendy Alvey and Beth Kilss, for their editorial assistance and preparation of visual

aids for the presentation of this paper; Clementine Brittain, for her assistance with both the presentation and the paper; and Roselind Vinson, for her help in preparing visuals for the presentation.

REFERENCES

Barr, Richard and Turner, Scott J. (1978), "A New Linear Programming Approach to Micro Data File Merging," *1978 Compendium of Tax Research*, Washington, DC.: Department of the Treasury, Office of Tax Analysis, pp. 131-149.

Cilke, James M. and Wyscarver, Roy A. (1987), "The Treasury Individual Income Tax Simulation Model," *Compendium of Tax Research 1987*, Washington, D. C.: Department of the Treasury, Office of Tax Analysis.

Cilke, James M. and Wyscarver, Roy A. (1990), "The Treasury Individual Income Tax Simulation Model," *Compendium of Tax Research 1990*, Washington, D. C.: Department of the Treasury, Office of Tax Analysis.

Czajka, John (1988), "Development of a New Income Classifier for a Sample of Individual Tax Returns," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Czajka, John and Walker, Bonnye (1989), "Combining Panel and Cross-Sectional Selection in an Annual Sample of Tax Returns," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Czajka, John and Schirm, Allen (1990), "Overlapping Membership in Annual Samples of Individual Tax Returns," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Czajka, John and Schirm, Allen (1991), "Cross-Sectional Weighting of Combined Panel and Cross-Sectional Observations," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Deming, W. Edwards (1986), *Out of the Crisis*, Cambridge, MA.: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Dippo, Cathryn and Herrmann, Douglas (1991), "The Bureau of Labor Statistics' Collection Procedures Research Laboratory: Accomplishments and Future Directions," *Statistical Policy Working Paper 20: Seminar on Quality of Federal Data*, Washington, D.C.: Office of Management and Budget.

Hostetter, Susan (1987), "Measuring Income for Developing and Reviewing Individual Tax Law Changes: Exploration of Alternative Income Concepts," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Hostetter, Susan et al. (1990), "Choosing the Appropriate Income Classifier for Economic Tax Modeling," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Kennickell, Arthur B. (1991), "Imputation of the 1989 Survey of Consumer Finance: Stochastic Relaxation and Multiple Imputation," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Juran, Joseph M. (1988), *Juran on Planning Quality*, New York, NY.: The Free Press.

Lent, Janice (1991), "Variance Estimation for Current Population Survey Small Area Labor Force Estimates," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Nelson, Susan C. (1986), "Family Economic Income and Other Income Concepts Used in Analyzing Tax Reform," *Compendium of Tax Research, 1986*, Washington, DC.: Department of Treasury, Office of Tax Analysis.

O'Conor, K., Atrostic, B. K., and Gillette, R. (1990), "Moving From Descriptive Statistics to Inference," *Proceedings of the Symposium 90: Measurement and Improvement of Data Quality*, Ottawa, Ontario: Statistics Canada.

Schirm, Allen and Czajka, John (1990) "Intertemporal Stability in Total Income and the Overlap in Annual Samples of Tax Returns," *Proceedings, Section on Survey Research Methods*, American Statistical Association.

Schirm, Allen and Czajka, John (1991) "Alternative Designs for a Cross-Sectional Sample of Individual Tax Returns: The Old and the New," *Proceedings, Section on Survey Research Methods*, American Statistical Association.