STATISTICS FROM INDIVIDUAL INCOME TAX RETURNS: QUALITY ISSUES

Beth Kilss and Fritz Scheuren, Internal Revenue Service

One of the main reasons for the rapid growth in the statistical uses of administrative records during the 1970's was that certain data could be obtained at a relatively low cost and without increasing respondent burden [54]. Cost and respondent burden are likely to continue to be important factors in the statistical use of administrative records in the current decade given the recent changes in the Federal government's statistical programs [39]. Even though administrative records are being used statistically more and more all the time, there are very few examples of administrative record systems that have been designed with statistical uses in mind. Instead, statisticians, economists and researchers have had to locate existing administrative data that best suit their intended use. In other words, most statistical uses have been developed in an ad hoc manner. The statistician is not in control of the design and collection of the records so as to yield desired statistical characteristics. Consequently there are many unanswered questions about the quality of Federal statistics from administrative records.

This paper discusses the quality of information obtained from one of the Federal statistical data systems derived from administrative records--the Internal Revenue Service's Statistics of Income (SOI) program for individual income tax returns. The main theme is that quality concerns are growing--paradoxically at the same time as reliance on administrative records for statistical purposes has been increasing.

The material in the paper is divided into three main parts. The first of these is a brief analysis of the SOI program. This case study illustrates and quantifies a number of dimensions of what is meant by quality. It also makes concrete the usual processing and reporting environment in which statistics from administrative records are developed.

In the second part of the paper we discuss the impact on quality of dollar budget and burden budget cuts. Quite obviously, with the recent cuts in the statistical budgets, virtually all the statistical agencies and non-statistical agencies are threatened with a serious "attack" on the quality of statistics, whether from surveys or administrative records. The key for each agency is, of course, to deal with these cuts so as to minimize, to the greatest extent possible, their impact on final delivered quality.

The paper concludes with some observations about the kinds of issues that warrant consideration if the statistical community is to preserve the needed quality of Federal statistics given the likelihood of still further dollar budget and burden budget cuts.

1. STATISTICS OF INCOME (SOI) PROGRAM

Shortly after the passage of the constitutional amendment calling for the reinstitution of the income tax, IRS began tabulating information from tax returns in its Statistics of Income program. (There was a Federal income tax during and shortly after the Civil War, which was declared unconstitutional. The income tax was reinstated in 1913 with the ratification of the Sixteenth Amendment to the Constitution and the enactment of the Revenue Act of 1913.)

One part of the Statistics of Income program involves collecting and processing data from the individual income tax returns. The program for these data has developed over the years and, now, microdata files for public release are produced annually by the SOI Division at the Internal Revenue Service. Statistical information is also tabulated and published in a volume of the Statistics of Income series [59]. The statistics produced in the SOI program for individuals are based on a stratified probability sample of unaudited individual income tax returns and represent coverage of the Forms 1040 and 1040A filed by U.S. citizens and residents for a particular income year. For Income Year 1979, the SOI sample consisted of about 200,000 returns taken from the 92.7 million that were filed [59].

Responsibility for the development of the tax forms for the coming filing season lies with the Tax Forms Committee who coordinates the activity with the Joint Committee on Taxation, with other interested members of the Treasury Department, as well as with the United States Congress. Internal Revenue Service provisions needed because of tax law changes and attempts at simplification are also coordinated with the Tax Forms Committee. In addition, there is now an involvement with the Office of Management and Budget (OMB) in this process as a result of the Paperwork Reduction Act, which requires that OMB clear all tax forms.

In a normal year, the individual income tax forms are finalized in the summer and sent to the printer in September or October. In 1981, the process was compressed somewhat and complicated because of the Economic Recovery Act which was not signed until August. Generally copies are available in December from banks and from other public places for people who did not file in the previous year, or who had moved and for whom the "tax package" might not be forwarded in time. In January, the Postal Service sends copies to individuals who filed the previous year.

The SOI Division plays a minor role in the development of the tax forms. Usually, information about the usage of prior tax forms is provided by the Division to the Tax Forms Committee. The Division also plans its statistical program around the proposed tax forms. There is extensive consultation with SOI users, notably the Office of Tax Analysis, learning what information the users are expecting to obtain from the returns. Occasionally, the Division will request tax form

changes to simplify its work or to meet user requirements. In any event, it cannot complete the planning of its programs until the final forms have been decided. Nonetheless, during the summer and fall while the forms are getting into shape, the SOI Division is developing the statistical system which will process these forms.

Development of the computer systems for statistical purposes

About the same time the return forms are finalized, the development of processing instructions and training materials to be used in the processing of these documents begins. It might be noted in passing that surveys have a distinct advantage over administrative records regarding the development of computer systems for statistical purposes. In a survey, like the Current Population Survey (CPS), for example, the control of the questionnaire is more or less completely with the statistical agencies that are employing the form. In the case of SOI, that is not true, which increases the costs considerably. Statistics of Income undergoes a substantial redesign every time there is major tax legislation. (For a brief summary of major tax law changes affecting individual income returns from 1917-1979 see [6]).

Tax law changes have an impact on quality, of course, in several ways. One is that systemic problems arise and are embedded in statistical systems. If you are running essentially the same system for a long time, you eventually may uncover nearly all of the problems that exist in that system. If you continuously revise a system, you have to "re-system" test it each year and the chances of actually uncovering the errors that exist, the mistakes that are systemic, are less in such an environment. Obviously, therefore, some element of stability in a statistical system is a key to quality. It is hard for an agency which is producing statistics from changeable administrative records to achieve stability. To do that requires a different kind of ingenuity than at least we have been able to fully implement so far. The newer computer technologies which allow for generalized systems and the degree to which we can apply them will be a test of to what extent we can actually address these underlying systemic quality issues.

In talking about systems as a whole, there is an observation that can be made about quality. In an environment where quality is being measured, typically 85 percent of the failures in quality are systemic; only 15 percent are due to individual quirks or individual coding, etc., errors [16]. Errors either get into the system or stay in the system because of a systemic flaw. In an administrative environment where the form is changing outside of the control of the statistical unit, these systemic flaws are very much harder to detect and protect against because so much of the available resources have to be put into revising those systems year in and year out. In any event, there are a whole host of steps that the SOI Division takes to develop the statistical system that it is going to employ for a particular tax year's forms, and given that year's particular set of instructions. These include the development of sample design criteria, computer software, processing instructions, clerical training materials and so forth.

The data capture process

The administrative data capture and the supplementary SOI statistical data capture processes work hand in hand. Each one has quality checks embedded in it. First of all, the Form 1040/1040A is filled out by a taxpayer. Since not all taxpayers comply with the filing requirements, there is a coverage error in the data. Some measurement of this has been done. For 1976, the estimate was that about six million, or a little over 8 percent of the total eligible tax filers actually failed to file returns [58]. Most of those "taxpayers" were ones that, in fact, had they filed, would have received a refund, since the wage withholding system had collected more tax than these people had to pay. There were, of course, flagrant abusers of the system who, in fact, owed considerable sums of money; but, typically, in terms of the number of taxpayers, most of the taxpayers in fact could have had a refund if they had filed.

There are "coverage improvement" activities underway routinely, but since Statistics of Income is based on pre-audit information, these coverage improvements in the current year do not have an impact on the current year's SOI; coverage improvements in prior years do have an impact in terms of improved taxpayer performance.

When the return is filed, it is processed in one of the ten IRS service centers around the country. (People in Washington, DC, for example, mail their returns to the Philadelphia Service Center.) The mail is opened and if there is a payment due, the checks are immediately separated from the returns and cashed. The returns are then examined for completeness and passed on to be transcribed and key-entered onto a disk and then onto tapes, called "transaction tapes," containing virtually all of the information on the main part of the return, and much of what is on the schedules that have to be attached. These transaction tapes are sent to the National Computer Center in Martinsburg, West Virginia, where they are, as the phrase is, "bumped up" against the Individual Master File (IMF) of all taxpayers. During this processing, certain information is posted to the IMF from the transaction tapes and comparisons are made with what taxpayers did in former years.

It is at this point that the sampling of the tax return information is done for statistical purposes, generating another computer tape to pull the returns which have been filed in the service centers after the initial administrative data capture. For the most part, historically the samples have been reasonably efficient for the many purposes for which they were intended despite the fact that

last minute law changes can make the sample designs out-of-date.

SOI users typically have not thought the samples large enough, however, even though in the 1040 area they have been running about 170,000 returns or more (there were 203,605 returns for 1979). Because of the recent budgetary constraints, the SOI Division has been forced to cut sample sizes. They have been reduced to about 135,000 returns for Income Year 1981 (now being processed). Plans for 1982 call for only about 115,000 returns. Techniques to maintain the reliability of the key statistical series are being instituted to compensate for these cuts, to the extent possible. For example, a greater use of longitudinal samples is among the methods being considered to preserve the reliability of estimates of year-to-year change. The quality of the overall cross-section statistics will in part be preserved with the increased post-stratification to Master File totals [6].

Administrative quality improvements

Before describing the statistical part of this system in more detail, the nature of the quality checks that are imposed on the administrative system will be discussed briefly. When returns are filed initially with the ten service centers, they are processed for administrative purposes to determine the correct tax liability. As part of the key-entry process, there are a whole set of procedures known as "consistency tests" which are employed to identify the simple mistakes; i.e., mathematical errors made by taxpayers and mistakes made in the actual data capture process itself. (Consistency tests are tests as simple as adding up all income to see if it adds up to the total, calculating or recalculating the tax based on information from the return, and looking for questionable or unusually large or small reports for particular items.) There is also a service center quality review program to measure and limit errors in administrative processing such as coding, editing and transcription.

Later on in the administrative processing, statistical procedures are employed to classify taxpayers, using discriminant function techniques, into those that are likely to yield large tax changes--either positive or negative. The Internal Revenue Service's mission is not to collect more taxes; it is to see to it that people pay their proper tax. Thus, an effort is made to find taxpayers who have considerably overpaid their taxes as well as taxpayers who have, of course, underpaid. In any event, the system does classify taxpayers in terms of "audit potential," as the phrase is, but that is not done at the initial data capture step.

Checks at the initial data capture stage are confined largely to the general sorts of mathematical errors, keying errors, and copying errors that occur in any large system. (See figure A below.) In terms of mathematical errors made by taxpayers, two out of three of the over six million detected mathematical errors for 1979 were in favor of the taxpayer. This tendency for mathematical errors to favor the taxpayer has been such a longstanding historical phenomenon that perhaps we should just make an observation about it.

The first thing to dispel is the notion that the persistent direction of the errors made each year indicates that there is a wholesale attempt to bilk the government. On the contrary, most people, as a patriotic duty, try to pay their fair share; what may be happening is that there is a sense of frustration that comes into play when we calculate our tax (as we sit there in the "dark of night," typically pictured in the cartoons, with a candle lit and a lot of scrap paper and pencils--several broken ones, perhaps--anyway, the erasers completely used up). During these calculations, we find an answer which is "reasonably small" and we think it must be right, and so we end up with that figure. If the number seems too high to us, we recalculate it again. And so there is, in a sense, a stopping-rule problem here. Taxpayers tend to stop when they get the answer they want, which is not always the right answer. And it is not a conscious attempt to cheat; it is simply part of the process of reporting on one's income.

Taxpayers do, as the conventional wisdom asserts, tend to underpay their taxes in general, rather than overpay them. There is considerable evidence, some compiled by IRS as part of its Taxpayer Compliance Measurement

Figure A.--Taxpayer Error Rates, Tax Year (TY) 1979

Types of Errors (TY 79)	Percent of All Errors Made
Math errors	32.0
Wrong line entries	19.0
Earned income credit errors	16.0
Incorrect (illegible) entries	11.0
Wrong tax table/tax rate schedule	7.0
Withholding errors	5.0
Other credit errors	4.0
Miscellaneous errors	6.0

NOTE: Unsigned returns and missing/incorrect social security numbers are not counted as errors in this case as returns with these deficiencies are not processed until corrected. Data based on a special compilation done at the Ogden Service Center. While strictly applicable only to taxpayers who file in Ogden, the data probably reflect accurately the general nature of the errors made by all taxpayers for 1979.

Program, some available by comparison with other statistical series, such as those that are developed by the Bureau of Economic Analysis, suggesting that in fact there is a systematic underreporting of some income types and consequently, underpayment of taxes. The impact of this factor on quality is documented in a number of places. (See for example [58].) Developing a strategy for making this kind of information available to users of Statistics of Income data so that it is relevant to their use of SOI statistics is a very hard problem. Certainly, it is a problem that has not been given enough attention for us to report on it here [22], beyond noting the major difficulty that is imposed by trying to integrate little "snippets" of information from a lot of different sources and some major pieces of information from within IRS in a context of converting pre-audit and pre-examination results from Statistics of Income into post-audit, postexamination results.

Mathematical errors corrected during administrative processing have an impact on the data that are included in Statistics of Income. The elimination of these errors, needed to ensure the accuracy of the information for administrative purposes, is also a quality improvement step for the statistical purposes of the SOI program. However, not all information collected on the individual income tax return is of equal importance for tax administration versus statistical purposes. Resources needed to ensure accurate reporting and processing of information will be concentrated where administrative priorities lie. Therefore, reporting and processing errors can have a negative impact on the accuracy of items which may be very important for statistical applications, but not so for tax adminstration.

Separate statistical operations

After the sampling at the National Computer Center, the selected records are sent back to the service centers on tape and the sample pulling begins. Additional errors can occur here, of course. For example, it may not be possible to find a particular return. In a large "paperwork factory" like a service center, returns are not always "locatable" (usually about 0.5% of the returns are unavailable for statistical processing). For one thing, they may have gone on to some other administrative function. To the extent possible, the return is copied or otherwise retained once it is selected, even though it may be needed almost immediately for an examination, or for some other taxpayer contact (for a continuing investigation, for example, in a fraud case). In any event, not all the returns which are in the SOI sample are found. Typically, there are two options in such cases. The return can be excluded from the sample and a unit nonresponse adjustment made. This option tends to increase the sampling error and can bias the results (but generally, unless the return has unusually large income amounts, the bias is small). The other alternative is to try to use the information which is already available from administrative records and "piece together" or impute the statistical information that is needed but not on the administrative file. In other words, make an item nonresponse adjustment which also will have bias and variance consequences. (This latter procedure is preferred for high-income taxpayers.)

For the bulk of the sample returns, additional information [64] is manually edited and transcribed for merging to the computer record obtained from the administrative processing.

Statistical testing in the service center is accomplished while the return is still available [29]. These tests are an important factor of the SOI program because certain kinds of errors are not uncovered in administrative processing as they do not impact on the administrative purpose. For example, a very low income taxpayer who may not have even needed to file a return might have a record with errors on it which did not get corrected because there were no tax consequences. These errors would impact on the statistics from the records, however, and so they need to be examined and an adjustment procedure developed for them.

When errors are found, the return is re-examined and an attempt is made to resolve the errors in the service centers. Some of the errors that are uncovered are taxpayer errors that are not necessarily resolvable using the return. In these cases, a judgment is needed that is based on the end use of the data, or at least the perceived end use. In order to control these judgments, and to carefully monitor how they are made, there is a later processing stage when all of the data on tape are shipped to the Detroit Data Center and combined and retested [65]. Final adjustments are made with the direct involvement of the subject-matter specialists who are representing the final users and who are responsible for managing the production of the tabulations.

Notwithstanding, at the service center, the error readout is fairly substantial, on the order of 16 to 18 percent in this past year [19]. Many of these errors are minor keying errors. Others are editing (abstracting) errors. A number of the nonkeying errors--well over half--are made by taxpayers. Obviously, some so-called "taxpayer errors" are in fact not errors at all, rather they are unusual circumstances which have been erroneously treated as errors. SOI consistency testing may thus introduce mistakes in the data themselves by changing the records in conformance with a model of reality that may not totally fit all situations. Statisticians seem to have an occupational vice; they tend to over-correct or over-smooth (in this context at the microdata level). This predilection, it goes without saying, can be very, very damaging if it is not carefully controlled because it forces on the end user the data model of the data producer and if the data model of the data producer is a very tight model requiring a great deal of consistency in the data relative to the model, then some of the information--maybe even the very thing that is most important--could ultimately be lost. (There is a philosophical

issue here about how much error you keep in and how much error you try to take out; someone's error may be someone else's information.) Of course, in addition to possibly "correcting" data that were correct to begin with, error resolution processing may incorrectly "correct" errors made in abstraction and keying.

As a way of controlling the introduction of error into the data due to "overcorrection," a second set of testing is done centrally in the Detroit Data Center for cases where professional judgment has to be brought to bear. Subsequent corrections, at least in terms of the manual adjustments, result typically in one percent of the data being examined or re-examined at this time. At this second round, a great many automatic edits and imputations are also introduced so there are minor impacts on nearly all the data. For example, in the tax system we only balance within tolerances in our administrative processing; for SOI the balancing is done to the dollar (imputing the out-of-balance portion).

It should be noted here that in recent years, there has been a noticeable decline in the proportion of sample return records that need to be corrected manually. Figure B shows the average error rate in Data Center consistency testing for an eight year period. For 1973, the error rate was a moderate 3.3 percent, but by 1980 it had declined to less than 1 percent.

Three major factors have contributed to these decreases. First, beginning with 1974, data items on the edit sheet were preprinted from the Individual Master File transaction tapes. The second major factor was the introduction for 1978 of an abbreviated Service Center error resolution while the returns were still available. A third factor has been a continuing expansion of the automated consistency tests, permitting the computer to correct an ever larger proportion of the errors in lieu of manual corrections [29].

Other quality assurance activities

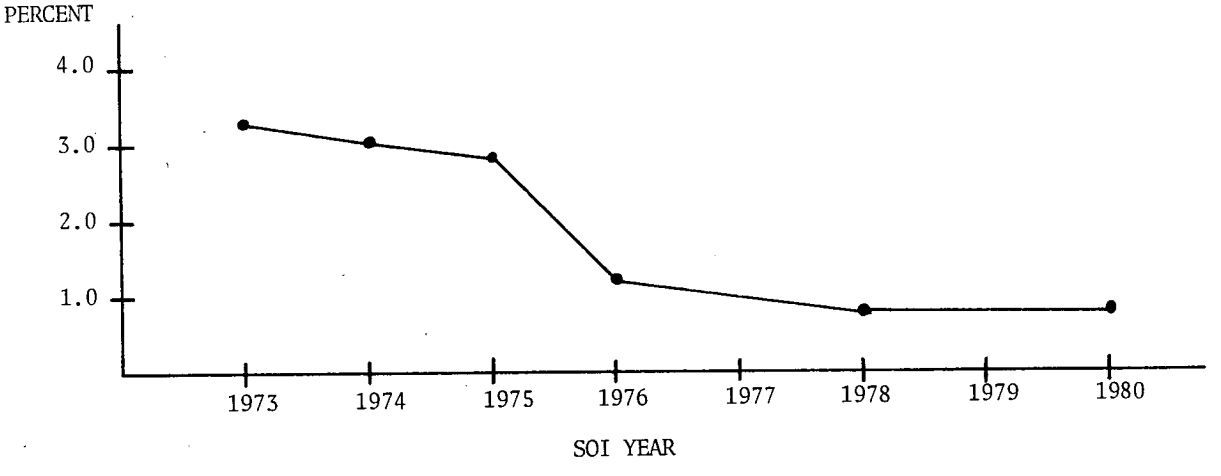
Errors, error checks and "corrections" occur at all stages in the processing of SOI. Some of the impacts on quality are measurable, some are not. In each stage of this system, therefore, there are additional quality review and quality improvement procedures--for example, testing is done of the system itself (systems acceptability testing) and also of the data that are flowing through that system.

In the manual editing (abstraction) of additional information, there is what has been called historically an "edit verification" step at which the manual transcriber's work is reviewed by another person on a sample basis using continuous sampling procedures, which depend on the individual transcriber's error rate (the people who make a small number of errors being reviewed less frequently than the people making a large number of errors).

The clerks at the beginning, of course, are all reviewed on a 100-percent basis until their work is sufficiently accurate to qualify for sample review. These procedures typically uncover an important number of errors, which vary by type of return. The simple returns, of course, may not have needed any review at all after the beginning stages of processing; in some of the more complicated high-income returns (returns with business schedules, and so forth) not all the errors will be uncovered even though they are reviewed on a 100-percent basis.

During key-entry, there is typically a 100-percent key-verification of the data, unlike the manual abstraction which is on a sample basis. This seems, of course, wasteful since much more time is spent reviewing correct data than resolving incorrect data. The SOI

Figure B.--Average Error Rate in Data Center Consistency Testing of Individual Income Tax Returns for Selected Years, 1973-1980



NOTE: Error rates are not shown for Tax Years 1977 and 1979 because the item content is not comparable to the earlier (and later) periods.

Division is examining this processing at this time. But it does unquestionably improve the quality of the data that are produced in SOI. While trying to save money in this budget environment, the Division is trying to find a way to preserve as much of that quality as possible. The already existing computer consistency tests procedures may work to detect most of the mistakes corrected by key-verification. The consistency testing packages though, however carefully they are developed, are not going to identify all of the data capture errors that are made; but nearly all of them may be identified and benefit/cost ratios would suggest that such key-verification could be eliminated.

The SOI Division has done an experiment in this area with partnership returns and derived interesting data concerning differences in error rates between systems which are key-verified and not key-verified. Basically, the empirical evidence seems to suggest that if you took the square root of the error made in a system in which key-verification has taken place, that would be about the error you would get if you did no key-verification [66]. This assumes that morale, motivation and "all of that" can be held constant. There is a concern, of course, that if you do no key-verification, people will get sloppy. In any event, some kind of a sample procedure seems to be an appropriate middle ground for preserving quality at roughly the same extent that now exists and, more importantly, measuring that quality--something that is not being done directly because of the nature of a 100-percent verification system, at least the one being employed, which does not permit us to routinely detect what the error rates are at initial entry.

In addition to obvious data capture errors that can be made, there are also errors which are introduced in the estimation side of the system, in the way the data are weighted. One of those was alluded to earlier, in terms of the handling of missing returns. Other problems, of course, have to do with the development of appropriate population totals and their introduction. Errors are made and checks have to be introduced in order to prevent such simple things as data reformatting of the microdata files, and poor table programming, from introducing mistakes into the tables.

Finally, errors of analysis can be made by the end users of the data. Here, of course, is what may be the most fundamental challenge to data producers. In order for the user to properly convert the data supplied to him into meaningful information, we need to provide better measures of quality that are related to his use and also more usable, more readable documentation of what was done and how it is going to affect the user's inference. There has been a general failure on the part of data producers to provide the latter.

By and large, the users do not tell the producer what quality they desire, other than an implied "error-free" product which may be impossible or too costly to achieve and in most instances is not needed for the use made of the data. Under these conditions, the producer often arbitrarily inserts quality measures into the various processing phases, too much in some, too little in others, with no definite goal and little, if any, knowledge of the quality of the finished product. At IRS, for example, many of our users for some applications might conceivably handle data having less quality than is being given them now, provided there were much better documentation of that quality.

Users of any Federal statistical data must regularly have access to reliable indications of the kinds and levels of error to which those data are subject. Such yardsticks are extremely important since without them users may rely excessively on statistics, presuming an accuracy much greater than actually exists. Without these indicators, officials in Congress and the Executive Branch will have to judge quality and other statistical improvements with no solid basis for weighing the benefits of added accuracy, for example, against added costs [22].

2. IMPACT OF DOLLAR BUDGET AND BURDEN BUDGET CUTS ON THE QUALITY OF FEDERAL STATISTICS

The Statistics of Income program, despite its long history and the changes to make it more cost effective and to improve its quality, is subject to many problems affecting statistics in general and statistics from administrative records in particular. Specifically, changes in the amount, kind and quality of statistical information that is collected are directly tied to changes in both the dollar and burden budgets. The implications of these for individual income tax programs will be developed in this section.

In 1981, there were approximately 38 statistical agencies, units, or programs, scattered through more than 90 Federal agencies, producing statistical data and information [8]. The growth in the role of the Federal government since the late 1940's has resulted, until recently, in an increase in the organizations and resources devoted to statistics and to the collection and analysis of governmental data. However, a report by the Congressional Research Service [39] found that between FY 81 and FY 83, the total budget of the Federal government's statistical programs will be reduced by 5.1 percent.

Dollar budget cuts

In March of this year, the House Subcommittee on Census and Population held a day-long hearing to examine the impact of the proposed Reagan budget on the collection and dissemination of statistical information by the Federal government. The testimonies of the witnesses represented the private sector response to Federal statistical data cutbacks. In particular, many of the testimonies focussed on the impact of the budget reductions on the utility and quality of Federal Statistics [e.g. 13, 63].

In written testimony to the Subcommittee, Thomas B. Jabine stated that the quality of Federal statistics is high and that for the

most part they are reliable and timely. He notes that now, however, "...the ability of the Federal statistical system to maintain this high standard of performance is seriously threatened." [30] Jabine points out that in the past a significant part of the Federal government's resources for statistical programs were allocated ..." to areas that are of primary importance to the quality of statistics: standards, methods research and development, evaluation and professional staff development." He further emphasized that over the long term, the recent budget reductions, however, are "...certain to result in a serious deterioration of the quality of statistics." Jabine attributes the probable future decline in the relevance and credibility of these data to the fact that "...disproportionately large cuts are occurring in the areas that will determine quality in the future." Some examples of these areas include cutbacks in quality control programs in various agencies, reductions in funds for validation studies, delays in update activities, postponement of revisions to the Standard Industrial Classification planned for 1982, and reduction of funds for the professional career development of employees of statistical agencies.

James T. Bonnen warned in his remarks before the Subcommittee [7] of the need for better coordination of policy and priorities in Federal statistics in a period of budget constraint. He feels that major and permanent damage will be done to the Nation's capacity for informed decisions because of mistakes that will be made in the 1982 and 1983 fiscal year budgets for statistical programs. In his view, this is highly likely because the capacity for national level coordination of statistical policy is inadequate.

It is probably fair to say that the Federal statistical system is still in a state of "shock" over what has happened; the representatives of the private sector quoted above are articulating the sentiments of many others throughout the system. Virtually all the statistical agencies and non-statistical agencies producing statistics, whether from surveys or administrative records, will most likely not be able to continue with their current methods for improving, measuring, or preserving the desired level of quality because of budget restrictions.

In an April 1982 [39] report requested by the Government Operations Committee, the Congressional Research Service analyzed the effects the budget reductions have had on the amount and quality of statistical data available from the Federal government. They pointed out that the budget policies will tend "... to create the most problems for agencies that already use the most efficient procedures because they must reduce the information they collect, process and release." Here the implication of the cuts is different for survey and administrative record statistical systems. Unlike survey systems, often administrative record systems are based on very large samples; sample designs typically are unsophisticated and do not make full use of modern sampling techniques. Often, too, in administrative record systems, again unlike survey systems, automatic editing and imputation are still in their infancy. One of the key strategies for those managing statistical programs using administrative records, therefore, if the quality is not to suffer, is to introduce changes (already made in the survey area) in such a way that the impacts of the dollar budget cuts can be minimized. It may be necessary, for example, to increase the use of sampling for quality control and use greater ingenuity in running the programs which need to be maintained. There may be a need to abandon some programs altogether. For others, a strategy of better quality measurement, but lower quality may be the only alternative. For statistical programs where the quality will have to be maintained or improved, lowering the cost of operations will be the only answer available.

At IRS, the Statistics of Income program has also felt the impact of the budget reductions. A continual erosion of the SOI funding base during the past ten years has resulted in a deterioration in the basic programs to the extent that they are no longer able to fully serve the needs of their major users--the Office of Tax Analysis (OTA), the Joint Committee on Taxation, and the Bureau of Economic Analysis (BEA)--even though every effort continues to be made to offset budget restrictions with increased efficiency in the processing of data.

Congressional budget submissions in support of SOI have steadily declined between FY 1973 and FY 1983. By FY 1981, severe program contractions had become necessary even though significant improvements in productivity had been realized. The latter included development of software to produce tabulations and introduction of electronic composition for the preparation of tables for publication. As a result of the FY 1982 budget reduction, sample sizes had to be reduced considerably, while many data elements were deleted and several programs postponed or eliminated.

To address this situation, the Commissioner of Internal Revenue and the Deputy Secretary of the Treasury have concluded that non-OTA statistical requests should be funded on a reimbursable basis to the maximum extent possible. For example BEA is going to be asked to fund one-third of its needs in FY 1983 and the full amount in FY 1984. At this time this funding issue is still unresolved.

Burden budget cuts

The general public is by now quite familiar with the current Administration's dollar budget restrictions; the impact of the Paperwork Reduction Act, however, which imposes "burden budget" cuts on all agencies of government is less well-known. Reductions of this kind are related to the amount of paperwork that each agency requires of the public.

The Federal government first sought to control the growth of reporting burden by the Federal Reports Act of 1942. Over the years, however, the clearance authority became very fragmented, undermining the capacity for coherent burden reduction. In recent years,

however, the problems of burden control have been addressed with the establishment of a "burden budget" process (by Executive Order No. 12174 of November 30, 1979), and the enactment of the Paperwork Reduction Act of 1980. The latter completely revised the Federal Reports Act and returned, as of April 1, 1981, all final clearance authority plus statistical policy and standards to the Office of Management and Budget [8]. It establishes a quantitative limit on the amount of paperwork that agencies may require of the private sector or State and local governments in collecting Federal data in order for the government to function. Sizeable cuts are mandated by law; by the end of Fiscal Year 1983 there is supposed to be a reduction in burden of about 25 percent [7].

The extent to which the government is able to manage Federal reporting requirements efficiently and wisely will have a direct impact on the amount and quality of the data produced by the statistical system. James T. Bonnen, in his testimony to the House Subcommittee on Census and Population last March, expressed fears that the burden budget on Federal statistics is "...distorting statistical priorities." He also stated that "...through Fiscal Year 1982 in most statistical agencies, the burden budget has had far more impact on statistical products than the dollar budget. It has also very likely distorted priorities between regulatory and other program agencies and statistical agency collections and uses [7]." A recent paper by the Congressional Research Service cited earlier [39] reported that "...OMB is preparing a new forms clearance guideline that will very likely give greater weight to establishing a Federal need for the data as distinct from broader national needs. Agencies that sought to justify data collection to meet the needs of States, local governments, or private organizations would be required to show that no alternative was available." There is a brief discussion in [8] of some of the ways in which the management of Federal reporting requirements and the statistical system interrelate. A summary of the main points may shed further light on what the burden budget cuts could mean.

First, the tools used by statistical agencies--sampling, quality control, analysis of existing data, etc.--are near the roots of reporting requirements and can reduce reporting burden if used appropriately. From the point of view of response burden, the use of appropriate statistical techniques is important, therefore, to all Federal data collection. Second, there are numerous major statistical programs which depend on the accuracy and quality of information compiled for administrative and other non-statistical purposes. Administrative records are used both directly (tax return statistics) and indirectly (survey frames for statistical inquiries) in statistical programs. Thus, the level and quality of responses to reporting requirements throughout the government has an important, although often indirect, bearing on the quality of statistical data.

What can be inferred here is that while the burden budget cuts will probably have an impact on the amount and quality of the statistics produced by the Federal government, the impact may be greater on statistics produced from administrative records than those from survey systems. There are a number of reasons that can be given for this. For one thing, the direct burden of the statistical data collection in surveys is very small relative to the total amount of Federal government paperwork (less than 1% in 1981).

On the other hand, cuts in burden can have an important impact on administrative record data. Administrative records systems typically require 100-percent reporting (which is then sampled for statistical purposes in many cases). The burden of the information required is, therefore, a great deal more initially, even though not all that information may be used for a statistical purpose. These cuts may affect the quality of the administrative data collection systems because burden reductions may limit the scope of the data items collected. Careful consideration should be given to the burden reductions so that important data items are not omitted or compromised to the point that user needs cannot be met. Agencies that sponsor administrative data collection must understand the needs of the ultimate user, as well as their own needs, in order to establish proper priorities.

There is a fear, shared by many, that in the course of reducing the administrative paperwork burden, the statistical needs of the society, the need for information, will not be given adequate weight in the decision process. Nor, when in fact it is agreed that certain information should not be collected on a 100-percent basis, will the ingenuity of the statisticians, which is quite impressive in the statistical system as a whole, be brought to bear on devising other methods of collecting that same information on a less than 100-percent basis.

In other words, a way must be found to reduce the burden on the general public; at the same time, the statisticians will have to obtain, possibly through sample methods imbedded within administrative records systems, the information that the society needs. Obviously, if information is not required for an administrative deliberation concerning a particular individual, an economic unit, or a reporting unit, it should not have been collected on a 100-percent basis to begin with. On the other hand, it may need to be collected on a sample basis. The model in the decennial census, of a limited 100-percent enumeration coupled with a more extensive sample data collection, is a model that at least in its general structure needs to be considered as part of the approach to burden budget reductions.

3. OTHER QUALITY ISSUES

There is no question in any of our minds that this is a critical period for the Federal statistical system as a whole. With limited budgets for data preparation and analysis, the coordination of statistical activities with respect to data production in terms of gather-

ing, analyzing, reporting and disseminating and data use will be increasingly important as scarce resources must be distributed among various statistical programs for their maintenance and continued development. Better cooperation is essential both among data producers and between data users and data producers.

To maintain or improve data quality the statistical units or agencies definitely should consider using the challenge of the recent dollar budget cuts as an "opportunity" to rethink ways to share information, and to move towards integrating their systems in areas of mutual concern. Obviously, the present lack of a strong coordinating unit at OMB will make progress in this area difficult. Nevertheless, better cooperation coupled with a focus on more creative approaches to Federal statistical programs can lead to a better integrated system of data collection and may lead to significant savings of resources both for the public in terms of reduced burden and for the government as a whole in terms of reduced cost. A recent proposal for exchange of industry codes between IRS and the Social Security Administration (SSA) by Thomas B. Jabine and Linda Bouchard Taylor [31] exemplifies the kind of ingenuity that may lead us to a resolution of both dollar and burden budget problems without materially sacrificing quality.

Industrial classification of identical or overlapping populations of economic units is carried out by a number of different agencies in the Federal government. A substantial portion of the volume of industry coding takes place at IRS, SSA, the Bureau of Labor Statistics, and the Census Bureau. This has led to problems of comparability for researchers and other users who try to combine or compare data from different agency sources. (A comprehensive review of industry coding systems, in fact, is the subject of another paper in this session [20] by Michael G. Farrell, Thomas B. Jabine and Carl A. Konschnik.) A look at the IRS and SSA systems in the case study by Jabine and Taylor shows that many of the same units are being coded by both agencies, and, therefore, there is much duplication of effort. The code-sharing proposal to reduce or eliminate this duplication will require further study, of course, but could result in advantages to both agencies.

As the budget process continues to unfold and limited resources become more of a problem, decisions about quality issues will have to be made with a much stronger focus on the user. Data producers must continue to meet the objectives of their statistical programs that are mandated by law, while at the same time satisfying the other needs of their users. The SOI program, for example, is the mandated responsibility of the Internal Revenue Service under section 6108 of the Internal Revenue Code. The Office of Tax Analysis and the Joint Committee on Taxation use the data to analyze existing and proposed tax legislation and to help estimate future tax revenue. The Bureau of Economic Analysis uses SOI as one of the essential building blocks in preparing estimates of the national income and product.

It is incumbent on the producer to focus more carefully on user needs. User funding is one strategy being explored at IRS so that user needs can be given a sharper focus. For the Bureau of Economic Analysis, this change will put in BEA's direct control decisions about the quality they want.

Challenges to the quality of the Federal statistical data base will continue in a climate beset by reduced funding and severely reduced, if any, central coordination. This should be reason enough, therefore, to work on methods to prevent any further erosion to the quality of Federal statistical information and to find more effective ways of communicating with users concerning their needs for data.

ACKNOWLEDGMENTS

The authors would like to thank Maria Gonzalez and Otto Schwartz for their contributions to this paper. Appreciation is also extended to Bob Wilson for his helpful comments, and to Joyce Coleman, Denise Herbert, Florine Morrison, and, especially, Nancy Robinson for their typing assistance.

NOTES AND REFERENCES

[1] Alexander, L. "Statistical Progeny of Administrative Records: Some Legal Issues." 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[2] Alexander, L. and Jabine, T. "Access to Social Security Microdata Files for Research and Statistical Purposes." Social Security Bulletin. August 1978.

[3] Alvey, Wendy and Aziz, Faye. "Quality of Mortality Reporting in SSA Linked Data: Some Preliminary Descriptive Results." 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[4] Aziz, Faye and Buckler, Warren. "Mortality and the Continuous Work History Sample." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[5] Bailar, Barbara A. and Kallek, Shirley. "Evaluation of the 1977 Economic Censuses." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[6] Blacksin, Jack and Plowden, Ray. "Statistics of Income for Individuals: A Historical Perspective." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[7] Bonnen, James T. "The Need for Better Coordination of Policy and Priorities in Federal Statistics." Testimony before the Subcommittee on Census and Population, Committee on Post Office and Civil Service,

House of Representatives, U.S. Congress, March 16, 1982. See also Bonnen, James T. "Federal Statistical Coordination Today: A Disaster or a Disgrace?" 1982 American Statistical Association Proceedings, Social Statistics Section,.

[8] Bonnen, James T., Clemence, Theodore G., Fellegi, Ivan P., Jabine, Thomas B., Kutscher, Ronald E., Roberson, Larry K., and Waite, Charles A. "Improving the Federal Statistical System: Issues and Options." Statistical Reporter, February 1981, Number 81-5, pp. 133-221.

[9] Bonnen, James T., Clemence, Theodore G., Fellegi, Ivan P., Jabine, Thomas B., Kutscher, Ronald E., Roberson, Larry K., and Waite, Charles A. "Improving the Federal Statistical System: Report of the President's Reorganization Project for the Federal Statistical System." The American Statistician. November 1981, Volume 35, No. 4, p. 185.

[10] Buckler, W. and Smith, C. "The Continuous Work History Sample: Description and Contents." Policy Analysis with Social Security Research Files. Social Security Administration. 1978.

[11] Cartwright, D. "Major Geographic Limitations for CWHS Files and Prospects for Improvement." Review of Public Data Use. March 1979.

[12] Cartwright, David W. and Armknecht, Paul A. "Statistical Uses of Administrative Records." 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[13] Casson, John J. "The Deteriorating Quality, Quantity and Utility of Federal Statistics." Testimony before the Subcommittee on Census and Population, Committee on Post Office and Civil Service, House of Representatives, U.S. Congress, March 16, 1982.

[14] Corby, Carol, Farrell, Michael, Blum, Alan, and Clark, Donald. "Content Evaluation of the Economic Censuses." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[15] DelBene, Linda and Aziz, Faye. "Mortality Coverage in Social Security's Earnings and Benefit Record Systems." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[16] Deutch, Jacob. "Design and Methodology of the Social Security Administration's Evaluation and Measurement Systems." 1976 American Statistical Association Proceedings, Social Statistics Section, Part I, p. 8.

[17] Dumais, James and Shadid, Raymond. "Individual Statistics of Income: Advancing the Closeout Date." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[18] Duncan, J., and Shelton, W. Revolution in United States Government Statistics 1926-1976. U.S. Government Printing Office, Washington, D.C., 1978.

[19] Durkin, Thomas M. and Schwartz, Otto. "The SOI Quality Control Program." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[20] Farrell, Michael G., Jabine, Thomas B., and Konschnik, Carl A. "A Review of Industry Coding Systems," 1982 American Statistical Association Proceedings, Section on Survey Research Methods.

[21] Farrell, Michael G. and Sullivan, John M. "Industrial Directory--Status and Direction." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[22] Federal Statistics, Report of the President's Commission, Volumes I and II, 1971.

[23] Fienberg, Stephen E. "Federal Government Statistics and Statistical Policy," Testimony before the Subcommittee on Legislation and National Security, Committee on Government Operations, House of Representatives, U.S. Congress, June 3, 1982.

[24] Freiden, A. "Data Development for Statistics of Income." Unpublished Statistics Division staff paper. An extract is included in the Appendix to "Using Business Master File Data for Statistics of Income Purposes," by William T. Powell and Joel R. Stubbs, Statistics of Income and Related Administrative Record Research, Internal Revenue Service, October 1981.

[25] Garnick, Daniel H. and Gonzalez, Maria E. "Statistical Uses of Administrative Records: Where Do We Go From Here?" 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[26] Gonzalez, Maria E., Ogus, Jack L., Shapiro, Gary and Tepping, Benjamin J. "Standards for Discussion and Presentation of Errors in Survey and Census Data," Journal of the American Statistical Association, September 1975, Volume 70, Number 351, Part II.

[27] Hanczaryk, Paul S. and Sullivan, John M. "Evaluation of Coverage of the Administrative Records Frame for the 1977 Economic Censuses -- Employer Segment." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[28] Hirschberg, David W. and Renshaw, Vernon. "Access to Administrative Records on Establishments and Individuals for Public

Policy Analysis." 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[29] Internal Revenue Service, Office of Assistant Commissioner (Planning and Research). "Current Mission and Operations of the Statistics Division." Volumes I and IV. Summer 1980.

[30] Jabine, Thomas B. Letter to the Honorable Robert Garcia. Subcommittee on Census and Population. Committee on Post Office and Civil Service. House of Representatives. Washington, DC. March 8, 1982.

[31] Jabine, Thomas and Taylor, Linda Bouchard, "Sharing Industry Codes: A Case Study." 1981.

[32] Judge, John F., Jr. "An Evaluation of the Geographic Coding in the 1977 Economic Censuses: An Overview." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[33] Kilss, Beth, Scheuren, Fritz and Buckler, Warren. "Goals and Plans for a Linked Administrative Statistical Sample." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[34] King, Carol S. and Ricketts, Edward K. "Evaluation of the Use of Administrative Record Data in the Economic Censuses." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[35] Knott, Joseph J. "Major Administrative Record Files: Documentation and Potential Uses." 1979 American Statistical Association Proceedings, Section on Survey Methods.

[36] Koteen, G. and Grayson, P. "Quality of Occupation Information on Tax Returns." 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[37] Leclerc, G. "Tax Administration and Economic Statistics: Some Conceptual and Policy Issues." 43rd Session of the International Statistical Institute. Presented in Buenos Aires, Argentina, December 9, 1981.

[38] Levine, Bruce. "Improving Industry and Place-of-Work Coding in the Continuous Work History Samples." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[39] Melnick, Daniel, et al. "Recent Changes in the Federal Government's Statistical Programs: An Overview of the President's Budget for FY 1983 and Analysis of the Departments of Energy, Labor and the Bureau of the Census." Congressional Research Service, The Library of Congress. April 8, 1982.

[40] Nathan, Gad and Sicron, Moshe. "The Use of Registers and Administrative Files for Preparing and Improving Statistical Data." 43rd Session of the International Statistical Institute. Presented in Buenos Aires, Argentina, December 9, 1981.

[41] Norris, Douglas, Leyes, John, and Kopustas, Nelson. "The Development of Social Data from Administrative Records: the Canadian Experience." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[42] Norris, Stephen. "Federal Government Statistics and Statistical Policy," Testimony for Governor Richard A. Snelling, Vermont, before the Subcommittee on Legislation and National Security, Committee on Government Operations, House of Representatives, U.S. Congress, June 3, 1982.

[43] Powell, William T. and Stubbs, Joel R. "Using Business Master File Data for Statistics of Income Purposes." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[44] Rosenberg, Harry M., Burnham, Drusilla, Spirtas, Robert, and Valdisera, Victor. "Occupation and Industry Information from the Death Certificate: Assessment of the Completeness of Reporting." 1979 American Statistical Association Proceedings, Section on Survey Research Methods.

[45] Rowebottom, L.E. "The Utilization of Administrative Records for Statistical Purposes." Survey Methodology. Volume 4, No. 1, pp. 1-15.

[46] Ryten, Jacob. "Error Detection in Statistics and Policies Governing Its Dissemination." 43rd Session of the International Statistical Institute. Presented in Buenos Aires, Argentina, December 4, 1981.

[47] Sailer, Peter J., Orcutt, Harriet, and Clark, Phil. "Coming Soon: Taxpayer Data Classified by Occupation." 1980 American Statistical Association Proceedings, Section on Survey Research Methods.

[48] Schwartz, Otto. "Quality Control in Statistics of Income: Past, Present, and Future." Unpublished Statistics Division staff paper.

[49] U.S. Department of Commerce. Bureau of Economic Analysis. Regional Work Force Characteristics and Migration Data (a Handbook on the Social Security Continuous Work History Sample and Its Application). 1976.

[50] U.S. Department of Commerce, Bureau of the Census. History of the 1977 Economic Censuses. December 1980.

[51] U.S. Department of Commerce. Bureau of the Census. Standard Statistical Establishment List Program. Technical Paper No. 44. January 1979.

[52] U.S. Department of Commerce. Office of Federal Statistical Policy and Standards. "Report on Statistical Disclosure and Disclosure-Avoidance Techniques." Statistical Policy Working Paper 2. 1978.

[53] U.S. Department of Commerce, Office of Federal Statistical Policy and Standards "An Error Profile: Employment as Measured by the Current Population Survey" Statistical Policy Working Paper 3 1978.

[54] U.S. Department of Commerce. Office of Federal Statistical Policy and Standards. "Report on Statistical Uses of Administrative Records." Statistical Policy Working Paper 6. 1980.

[55] U.S. Department of Health and Human Services. Social Security Administration. Economic and Demographic Statistics (selected papers given at the Annual Meeting of the American Statistical Association in Houston, Texas). November 1980.

[56] U.S. Department of Health, Education and Welfare. Social Security Administration. Policy Analysis with Social Security Research Files. Proceedings of a workshop held March 1978 at Williamsburg, Virginia, 1978.

[57] U.S. Department of Health, Education and Welfare. Social Security Administration. Statistical Uses of Administrative Records with Emphasis on Mortality and Disability Research (selected papers given at the 1979 Annual Meeting of the American Statistical Association in Washington, D.C.). October 1979.

[58] U.S. Department of the Treasury. Internal Revenue Service. Estimates of Income Unreported on Individual Income Tax Returns, Publication 1104 (9-79).

[59] U.S. Department of the Treasury. Internal Revenue Service. Statistics of Income. 1916-present.

[60] U.S. Department of the Treasury. Internal Revenue Service. Statistics of Income and Related Administrative Record Research (selected papers given at the 1981 Annual Meeting of the American Statistical Association in Detroit, Michigan, August 10-13, 1981). October 1981.

[61] U.S. Department of the Treasury. Internal Revenue Service. Statistics of Income Bulletin. Publication 1136. U.S. Government Printing Office, Washington, D.C., 1980.

[62] Wilson, Robert A. and DiPaolo, John. "Statistics of Income: An Overview." 1981 American Statistical Association Proceedings, Section on Survey Research Methods.

[63] Woolley, Donald E. "Budget Reductions and Federal Statistical Programs," Testimony before the Subcommittee on Census and Population, Committee on Post Office and Civil Service, House of Representatives, U.S. Congress, March 15, 1982.

[64] An example of an additional set of items might be the detailed calculation of the residential energy credit. Another illustration might be concerning the All Savers Certificate interest information, details of which might not all have been picked up, or not have been felt needed for tax administrative purposes, but are felt needed for tax policy examination of the new legislation.

[65] The Detroit Center is where all nontax processing is done at IRS, such as statistical work after the initial data capture, payroll recordkeeping, and so forth.

[66] In the experiment involving partnership returns, the error rate with verification was around 5 percent. The experiment was conducted without verification, and the error rate jumped to 20 percent. Since the square root of .05 is about .22, the result obtained should have been expected, given the square root model postulated. (The square root model for key-verification was suggested by Doyle Harris.)