

# STATISTICS OF INCOME QUALITY MANAGEMENT: INDIVIDUAL INCOME TAX RETURNS

Otto Schwartz and David Paris, Internal Revenue Service

This paper describes the means the Statistics of Income Division of the Internal Revenue Service is employing to foster better quality management in its programs and illustrates recent progress by looking at product quality results in the Statistics of Income program for individual income tax returns.

Organizationally, this paper is divided into several parts. The first of these provides a description of the Statistics of Income (SOI) program and historical quality control procedures. After that is a description of the current quality management procedures; then there is a report on the quality of the 1982 Individual Income Tax Returns SOI program. A brief mention of some of our plans for the future concludes the paper.

## BACKGROUND

Statistics of Income is an Internal Revenue Service program that collects, processes, analyzes and publishes information and data from tax returns. The data are needed and utilized by the Department of Treasury, Congress, Department of Commerce and various other government and private organizations for economic, financial and demographic research and analysis. SOI data are available from as far back as 1913. They cover all major tax returns, such as Individuals, Corporations and Partnerships, and a number of minor returns and forms such as Exempt Organizations and Private Foundations. The statistics are provided by various means, such as yearly SOI publications, articles in the quarterly SOI Bulletin, or microdata computer tape files.

Before going into the quality management aspects of the SOI program, a brief review as to how the data are derived may be helpful. The statistics gathered are determined by the needs of the data users. Figure 1 details the steps and functions involved in the production of SOI data. In summary, the tax statistics are abstracted from a sample of tax returns, keyed onto tape, computer-tested, tabulated and published.

Errors can occur on the tax return itself and in any of the mentioned processing functions. During the 1982 processing of individual income tax returns, IRS found errors on 8.9 percent of the returns or approximately 6 million arithmetic errors. (See also [1].)

Historically, SOI quality control procedures have consisted mainly of a systematic manual review and computer verification of the completed work at each processing phase [2]. The manual verification approach to manage quality had a number of drawbacks: chiefly good data were reviewed (causing inefficient use of resources); subsequent computer review, i.e., consistency testing followed by error resolution, was not considered (resulting in some duplicate review); lastly, despite these reviews at each stage, no overall measure of the

quality limitations of the SOI data could be provided to the data user [3], because the focus of the examination was on the processing and not on the final product.

## CURRENT QUALITY MANAGEMENT

Recently-imposed and continuing budget and burden (information reporting requirements) cuts have caused us to make numerous changes in the quality control aspects of the SOI program [4]. The SOI Division's current quality management approach consists of quality consideration in both the nonprocessing and the processing areas of the SOI program. Nonprocessing quality considerations include items that involve the "system" that provides the product. Processing quality management consists mainly of techniques to control, improve or measure the quality of the product or the effectiveness of the processing system. Specific quality management techniques utilized in each of the areas will now be described.

### Nonprocessing Quality Management Techniques.--

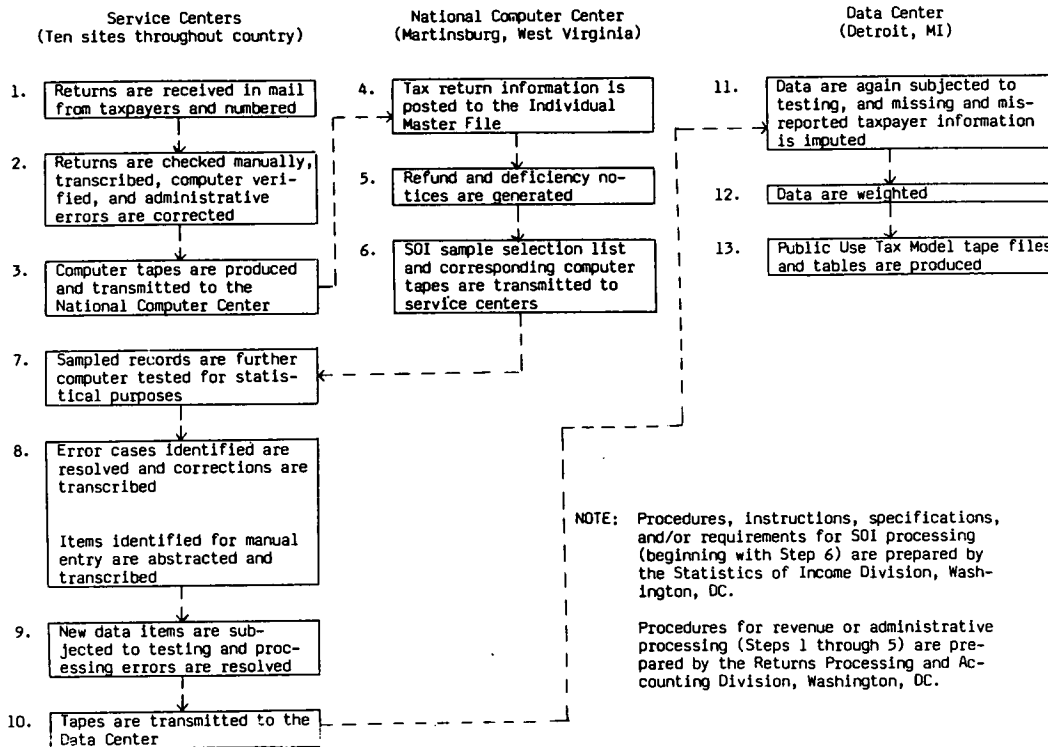
In the SOI Division, responsibility for quality rests with the branch that is responsible for the program. We think this enhances the prospect of an optimum quality approach as all functions of the processing are within the direct control of that branch and as the people responsible for quality are knowledgeable in the subject matter. This organizational structure furthermore allows a quick shift in the application of quality resources to assure their most efficient use and to address quality deficiencies in a timely manner.

Commitment to quality is stressed by SOI Division management. This commitment is expressed through the requirement of quality control procedures for all programs, encouragement of employees to obtain training in quality control and management techniques and the creation of a work atmosphere conducive to employee interest, participation and involvement in the division's activities. Recently, several courses in statistical quality control were conducted for SOI Division personnel and people in other areas of the IRS organization who are involved in SOI activities.

The addressing of human factors is of considerable importance. Division management is striving to increase the employee's involvement and participation in the organization and its programs, thereby increasing motivation and thus performance. Meritorious achievement by employees is rewarded. Employees are given the opportunity to select work schedules that best meet their needs as well as satisfy the needs of the organization. Employee attitude surveys concerning the effectiveness of the organization have been conducted and achievement of milestones in the organization's activities are celebrated.

The importance of effective communication can

Figure 1. SOI Data Source and Processing at Various Locations (1040 Returns)



not be overemphasized. The SOI Division employs various methods to foster and promote productive communication within the division itself and between the division and outside functions that are involved in the SOI activity. This is accomplished by such means as regular staff meetings; a written weekly highlights report indicating progress, developments and problems in the organization's activities; a weekly newsletter to and a weekly telephone conference call with field SOI personnel to discuss progress and problems and to inform; a yearly written multi-year operating plan defining goals both immediate and long-range; and an annual conference (face-to-face) with service center representatives for review and planning purposes involving all of the division programs.

The development and incorporation of standards both for quality and non-quality functions is being stressed. Some quality control experts, such as Dr. Deming, promote the philosophy of continuous improvement (thus, no standards) [5]. We feel, however, that our type of product can adequately satisfy user needs if they meet certain standards. Trying to surpass the standards may not be cost-effective. Also, since SOI data are produced in a multi-phase processing operation, it is sometimes more efficient to permit a high-productivity, relatively-high-error-rate phase if a subsequent processing phase (such as automatic computer correction) can economically correct the errors. A recent curtailment of 100-percent manual key verification with reliance on subsequent computer review is an example of this. Of course standards should not become

cast in concrete. This would fail to recognize that process improvement may be possible by using new equipment, better instructions, training or supervision, which can actually result in quality better than the standard at less cost.

Standards for non-quality functions such as scheduling, communication, productivity, and accountability are as important as product quality standards. Tight time frames for example can result in quality problems, as evidenced by an increase in the number of business returns with industry codes of "not allocable" as document processing productivity increases [6]. Lacking or improper accountability for actions can lead to a task not being done at all or at times being duplicated.

Effective use of available data is yet another nonprocessing quality management technique. We are living in an information age and we are being subjected to an information explosion. No data should be produced unless someone can make effective use of them.

Processing Quality Management Techniques.--The quality of a product and the process efficiency can be influenced by the quality management techniques applied prior to, during and after the processing. In SOI, our approach consists mainly of preparing written instructions that are as clear as possible, providing effective training, testing the computer systems that review data processing and prepare tabulations, processing a Preproduction sample [8] that tests the adequacy of the instructions and the effectiveness of the training, providing

computer assistance and review during processing, reviewing the tabulations and reports and processing a Quality Measurement sample which is described in the next section.

The quality of the processing is managed mostly through computer application. The computer tests the validity and consistency of the processed data, not only indicating possible errors for correction, but also directing the employee to enter certain data items. More and more of the computer review is being done on line; i.e., employees are given error information when the source document is still on their desks. However, some computer review is still performed after processing has been completed. Much of this involves automatic data correction or data imputation [9].

#### RESULTS OF QUALITY MANAGEMENT

It is said the proof of the pudding is in the eating. In SOI, the effectiveness of quality management is in the quality of the final data; e.g., our published reports and tape files. The quality of the final data (the computer file from which all tabulations are derived) is measured through a Quality Measurement (QM) sample. Until recently, this sample measured only the data abstraction function and thus did not portray subsequent data correction or adjustment, some of which may have been erroneous. The QM sample is one of our most useful and effective quality management tools. It not only provides quality limitations of the data when completed but also provides a rich storehouse of information for product quality improvement as well as process efficiency improvement.

Data reliability involves both sampling error (generally well quantified in SOI publications) and nonsampling (source data and processing) errors (generally discussed but not quantified). Nonsampling error should receive the same quantified treatment as sampling error currently does. This item is discussed further in the "FUTURE PLANS..." section of this paper. Errors should be quantified both in terms of numbers and dollar magnitude, for both affect the data use.

Following is a description of the QM sample, how it is processed, and the results of the review. All of the data are with respect to the 1982 Individual Income Tax Return SOI program (processed in 1983).

Quality Measurement (QM) Sample.---The QM sample is a stratified sample of returns from the SOI population. Stratification is by type of return. Figure 2 depicts the size and composition of the QM sample for 1982 and how it relates to the SOI sample and the U.S. population.

By the way, the relatively high proportion of nonbusiness-farm returns in the QM sample is due to the processing of an unscheduled additional seven thousand returns in this category at the specific request of the U.S. Department of Agriculture for a special study.

Figure 2.—SIZE AND COMPOSITION OF THE QUALITY MEASUREMENT SAMPLE

| Type of Return | U.S. Total (millions) | SOI Sample |                 | QM Sample |                |
|----------------|-----------------------|------------|-----------------|-----------|----------------|
|                |                       | Number     | Percent of U.S. | Number    | Percent of SOI |
| TOTAL          | 95.6                  | 238,277    | .923            | 943       | 1.068          |
| NB-NF ..       | 82.9                  | 34,055     | .041            | 98        | .289           |
| B .....        | 10.5                  | 44,091     | .419            | 251       | .569           |
| NB-F ....      | 2.2                   | 10,131     | .458            | 594       | 5.863          |

NB-NF = Nonbusiness - Nonfarm; B = Business  
NB-Farm = Nonbusiness-Farm

QM Sample Processing.---The QM sample review is conducted at the IRS Data Center in Detroit, Michigan, and consists of comparing the data on the computer document with the data on the corresponding original tax return. Also used in the review are all intermediate documents produced by the computer, such as error registers and corrected documents. An error in the review process is defined as an incorrect entry in the final computer file; i.e., the entry does not match the correct (or corrected) entry on the tax return.

QM Sample Review Summary.---It is commonly believed that a complex document will inevitably result in a high error rate. A 98.1 percent error rate found in a recent audit of hospital bills, for example, was not considered surprising [10]. The Form 1040 tax return is of course quite complex. The 1982 SOI program for Forms 1040 included, for example, about 600 potential items. Yet, as indicated in the following report, a high error rate does not necessarily ensue. Approximately ninety-six percent of the documents and 99.9 percent of the document entries (in the QM sample) were found to be error-free. Figure 3 presents these data weighted to the SOI sample and the U.S. total population. Data for 1981 are also included for comparative purposes. The differences between the two years are not statistically significant. Note that because of the relatively small sample size, large weighting factors, and technical constraints, data for less than total QM sample displays will sometimes not be weighted. When weighted, data will generally be displayed for the QM and SOI samples and for the U.S. Total. Variation due to QM and SOI sampling error is not included.

Figure 3.—OVERALL PERCENT ACCURACY OF SOI DATA

| Population Source    | Returns |      | Return Entries |       |
|----------------------|---------|------|----------------|-------|
|                      | 1982    | 1981 | 1982           | 1981  |
| Quality Measurement  |         |      |                |       |
| Sample .....         | 96.2    | 95.3 | 99.93          | 99.93 |
| Statistics of Income |         |      |                |       |
| Sample* .....        | 98.2    | 95.6 | 99.97          | 99.94 |
| U.S. Total* .....    | 99.7    | 97.3 | 99.99          | 99.97 |

\*Weighted

Also, although none of the defective QM sample returns were "forced" (the computer is forced to accept the document with a by-pass code at the completion of service center processing although it still has an "error" condition), the SOI file contains 2.5 percent such documents. The "error" condition or "inconsistency" in a forced return may in fact not really be an error, but an unusual though valid condition. The computer was instructed to impute data for forced returns to make them consistent. We are doing additional research on these returns to determine the quantity, nature and significance of possible error conditions.

Although the apparently high accuracy in overall SOI statistics (excluding forced returns) is very satisfying and might lead one to conclude that everything is in order and that things should continue as they are now, consideration of other-than-overall data displays and of process efficiency in terms of quantity/quality/timeliness/cost requires that we take another look at the data to determine how we can best benefit from them. Perhaps overall quality resources could be reduced, and, by reapportioning them, further improvements in quality may be possible in some not-so-reliable items. The subsequent discussion will attempt to address this issue by presenting error details by various breakdowns such as processing function, type of return, schedule and/or field on return, processing location, and magnitude of error.

Errors by Processing Function.--Processing errors for SOI data can occur in basically three functions: service center revenue processing, particularly transcription; service center SOI "edit" processing, including data abstraction, consistency testing and error resolution; and Data Center consistency testing and error resolution. Figure 4 shows that most of the errors still present in the accepted computer file are service center abstraction errors, followed next by service center revenue processing transcription errors and lastly Data Center error resolution errors. Returns with abstraction errors are the most common (1.1 percent of SOI returns), followed by those with service center revenue processing transcription and Data Center error resolution errors. The larger U.S. Total percentages in the transcription and error resolution functions are due to one error return with very large QM and SOI weighting factors.

Figure 4.—ERROR DISTRIBUTION BY PROCESSING FUNCTION

| Processing Function    | Percent of Entries in Error |      |      | Percent of Returns in Error |     |      |
|------------------------|-----------------------------|------|------|-----------------------------|-----|------|
|                        | QM                          | SOI  | U.S. | QM                          | SOI | U.S. |
| Abstraction....        | .05                         | .012 | .002 | 2.4                         | 1.1 | .14  |
| Transcription.....     | .02                         | .008 | .003 | 1.6                         | 0.8 | .18  |
| Error Resolution ..... | .01                         | .002 | .001 | 0.7                         | 0.3 | .16  |

Errors by Type of Return.--The processing errors are heavily concentrated in the nonbusiness-farm returns. This category involves about 89 percent of the defective QM returns, although it includes only 63 percent of the QM returns. Further, this category contains nearly 90 percent of the defective entries in the QM sample, or about 42 percent when weighted to SOI. The nonbusiness-farm category also has the highest error rate, both in terms of documents and document entries. When weighted to the U.S. Total, the difference in error rates among the various categories of returns becomes even more pronounced. See Figure 5.

Figure 5.—ERROR DISTRIBUTION AND PERCENTAGES BY TYPE OF RETURN

| Type of Return* | Percent of --   |                 |                 | Percent of --               |                  |
|-----------------|-----------------|-----------------|-----------------|-----------------------------|------------------|
|                 | Total Re- turns | Error Re- turns | Error En- tries | Returns in Error            | Entries in Error |
| QM Sample:      |                 |                 |                 |                             |                  |
| NB-F .....      | 63.0            | 88.6            | 89.1            | 5.1                         | .11              |
| B .....         | 26.6            | 8.6             | 7.9             | 1.6                         | .02              |
| NB-NF ....      | 10.4            | 2.8             | 3.0             | 1.0                         | .03              |
| SOI Sample:     |                 |                 |                 |                             |                  |
| NB-F .....      | 11.5            | 36.2            | 41.6            | (Same as in QM sample, with |                  |
| B .....         | 49.9            | 39.2            | 30.1            | sampling error)             |                  |
| NB-NF ....      | 38.6            | 24.5            | 28.2            |                             |                  |
| U.S. Total:     |                 |                 |                 |                             |                  |
| NB-F .....      | 2.3             | 48.8            | 44.0            | 6.5                         | .14              |
| B .....         | 11.0            | 44.2            | 51.7            | 1.2                         | .03              |
| NB-NF .....     | 86.7            | 7.0             | 4.3             | .03                         | .0004            |

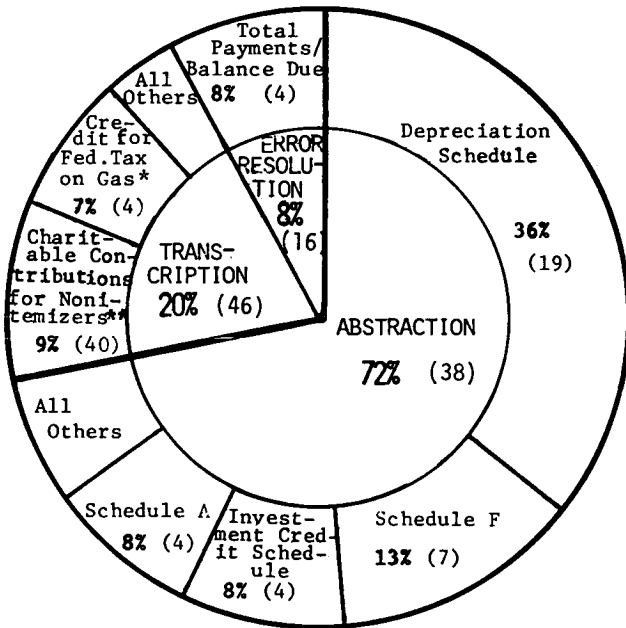
\*See explanation under Figure 2.

Errors by Schedule/Field of Return Form.--Errors occurred in 44 of the approximately 600 fields that had entries and averaged nearly three errors per return in error. Returns with two errors were by far the most common in the QM sample and again when weighted to SOI (an error in one field caused an error in another field), but when weighted to the U.S. Total, returns with four errors were most frequent (due to one return with four errors which had a very large weighting factor).

About one half of the QM abstraction errors involve the Depreciation Schedule (Form 4562). Another twenty percent involve Schedule F, Farm Income and Expenses, and about ten percent involve Computation of Investment Credit Form 3468 and Schedule A, Itemized Deductions. Errors in "charitable contributions for nonitemizers" and "credit for federal tax on gas" directly or indirectly account for almost all of the errors in the revenue processing transcription and Data Center error resolution functions. Figure 6 shows the most frequent errors by document schedule or item. Figure 7 shows all of the return fields with error rates exceeding one percent.

Errors by Service Centers.--Most of the editing errors occurred in three service

Figure 6.—ERROR DISTRIBUTION BY RETURN SCHEDULE/ITEM WITHIN PROCESSING FUNCTION (Weighted to U.S. Total)



\*This item caused all errors in error resolution.  
 \*\*Includes errors this item caused in Tax Table Income (4%) and Taxable Income (2%).  
 NOTE: Data exclude one error return with very large QM and SOI weighting factors. Numbers in parentheses indicate percentages of errors with this return included.

Figure 7.—RETURN FIELDS WITH THE LARGEST ERROR RATES

| Return Field Description                                 | Percent of Entries in Error |
|--|-----------------------------|
| Total Qualified Investment                               |                             |
| Nonrecovery Property .....                               | 5.39                        |
| Credit for Federal Tax on Gas .....                      | 4.55                        |
| Total Recovery Property Deduction for Current Year ..... | 1.84                        |
| Taxable All-Savers Interest .....                        | 1.70                        |
| Carryover of Unused Credits .....                        | 1.33                        |
| Business Net Income (Loss) .....                         | 1.12                        |

centers, slightly two-thirds. Transcription errors were also concentrated in just a few of the service centers, one center comprising slightly over three-fourths of these errors. Further, a single return was responsible for a substantial portion of the abstracting errors in three of the four service centers with the largest number of such errors. See Figure 8.

Errors by Magnitude (Dollar Amount).—The magnitude of an error is generally more important than the number of errors. In the auditing of hospital bills, for example, the patient finds it much more valuable to know that he is overpaying 4.9 percent per bill due to error than that, as

Figure 8.—ERROR PERCENTAGES AMONG SERVICE CENTERS BY PROCESSING FUNCTION (Weighted to U.S. Total)

| Service Center | Total     |           | Abstraction |           | Transcription |           |
|----------------|-----------|-----------|-------------|-----------|---------------|-----------|
|                | Re- turns | En- tries | Re- turns   | En- tries | Re- turns     | En- tries |
| Total ..       | .31       | .006      | .14         | .003      | .18           | .003      |
| A.....         | 1.76      | .036      | .38         | .009      | 1.38*         | .028      |
| B.....         | .56       | .009      | .46         | .007      | .20           | .002      |
| C.....         | .31       | .005      | .15         | .003      | .16           | .001      |
| D.....         | .24       | .004      | .10         | .001      | .14           | .002      |
| E.....         | .14       | .002      | .14         | .001      | .07           | .001      |
| F.....         | .10       | .002      | .10         | .002      | —             | —         |
| G.....         | .08       | .002      | .08         | .002      | —             | —         |
| H.....         | .05       | .001      | —           | —         | .05           | .001      |
| I.....         | —         | —         | —           | —         | —             | —         |
| J.....         | —         | —         | —           | —         | —             | —         |

\*One return with very large QM and SOI weighting factors comprises all transcription errors.

mentioned earlier, 98.1 percent of the bills are in error [10].

For most fields, the errors found in the QM sample generally had a negligible impact on the total dollar amount in the field in error. The dollar amount of the errors exceeded one percent of the total dollar amount of the field in error (weighted to the U.S. Total) for only three fields. See Figure 9.

Figure 9.—EFFECT OF ERRORS ON DOLLAR AMOUNTS IN RETURN FIELDS

| Dollar Change as Percent of Total Amount in Field | Number of Fields |
|---|------------------|
| Under .001 .....                                  | 4                |
| .001 under .01 .....                              | 2                |
| .01 under .09 .....                               | 22               |
| .1 under .9 .....                                 | 13               |
| 1.0 or more .....                                 | 3                |

Figure 10 lists the fields wherein the errors had the largest percentage impact on the total money amount. As a matter of interest, the 1.63 percent change in "total recovery property deduction for current year" signifies a change from the published \$8.5 billion to \$8.6 billion.

Errors by Type.—The errors were generally due to failure to enter (omission) or entering an incorrect amount, each type accounting for about the same proportion, about forty percent each for abstraction errors and fifty percent for transcription errors. The remaining abstracting errors were due to leaving incorrect entries, about fourteen percent, and changing correct entries, roughly five percent. The error resolution errors all involved changing correct entries, due chiefly to transcription omissions [11]. Field transposition and entry in the wrong field were rare in 1982 program but did pose a problem in 1981.

Omission errors generally appear to involve items that do not occur frequently on tax returns and thus the processing employee may not

be looking for them. This appeared to be the reason for most transcription omissions involving the fields "charitable contributions for nonitemizers" and "credit for federal tax on gas." In the abstraction process, most of the omissions involved depreciation schedule items.

Field transposition and entry into the wrong field appeared to be due to similarly-worded items that appeared close together on the tax return and/or on the abstraction document. Some incorrect transcription errors derived from transcribers who entered the first digit (or first few digits) of a multi-digit dollar amount and then, apparently distracted or interrupted, failed to transcribe the remainder.

Figure 10.--RETURN FIELDS WITH LARGEST DOLLAR ERROR IMPACT

| Return Field Description   | Percent Effect of Changes on Total Dollar Amount |
|--|--|
| Schedule D Capital Gain Distribution.....                                  | 1.63   |
| Total Recovery Property Deduction for Current Year...                      | 1.63   |
| Business Net Income (Loss)....   | 1.31   |
| Charitable Contributions (nonitemizers).....                               | 0.76   |
| Credit for Federal Tax on Gas.   | 0.67   |
| Computation of Investment Credit: Regular, New, Other, Basis Property..... | 0.59   |

Multi-Error Returns and their Significance.-- Although the average number of errors per defective QM sample return is 2.9, the number of errors range from 1 to 9. Three returns (containing 9, 8, and 6 errors, all editing) account for about 20 percent of the total QM sample errors. The errors involved mostly the Schedule F farm and depreciation schedule entries but were scattered among various fields within the schedule. Most multi-error returns involved an error in one (component) item which resulted in an interrelated (either "total" or an item calculated using the component or total value); e.g., an error in "total recovery property deduction for current year" involved an equivalent error in "total depreciation deduction."

Multi-Year Consideration of QM Data.--In order to reduce the relatively large sampling error due to the small QM sample, we have combined the 1981 and 1982 QM results for the changes in each field money amount as a percent of the total amount in the field. Figure 11 depicts selected common-to-both-years, relatively-high-error-rate fields as they appear combined and separate for 1982 SOI. Since errors tend to occur in different fields from year to year and the frequency of field usage tends to basically remain the same, combining two years' QM data generally halves the error rates.

Figure 11.--MULTI-YEAR ERROR DATA IMPACT ON AMOUNTS IN SELECTED RETURN FIELDS

| Field Description                           | Dollar Change Percent of Total Amount |      |
|---|---------------------------------------|------|
|   | 1981 & 1982                           | 1982 |
| Schedule D Capital Gain Distribution .....  | 0.94                                  | 1.63 |
| Business Net Income (Loss).                 | 0.66                                  | 1.31 |
| Credit for Federal Tax on Gas.....          | 0.34                                  | 0.67 |
| Current Year Regular Investment Credit..... | 0.17                                  | 0.34 |
| Form 4797 Basis.....                        | 0.13                                  | 0.22 |
| Taxable All-Savers Interest.....            | 0.10                                  | 0.19 |
| All-Savers Exclusion.....                   | 0.10                                  | 0.19 |
| Other Taxes Deducted.....                   | 0.10                                  | 0.19 |

FUTURE PLANS FOR SOI QUALITY MANAGEMENT

Utilization of Available Data.--We plan to make more and more use of available quality data to improve the quality of our product as well as of the processing system. Considering the quality assurance data for the 1982 SOI program, for example, we recognize that we need to "beef up" our SOI abstracting instructions and training, especially as concerns the nonbusiness-farm returns and accompanying schedule data, notably the depreciation schedule. A recently-redesigned depreciation schedule Form 4562 should help improve the quality of the data. In keying we have to find ways of detecting omissions and incomplete entries. We are currently in the process of obtaining contractor assistance to look into the quality problems in the keying process.

We will make the results of our quality measurement available to the data user to inform him of data quality limitations. This may take the form of referencing the QM report, including major findings, and/or making actual data adjustments within the published data.

Consideration of Non-SOI Functions on SOI Data Quality.--Since the SOI data are derived from a document which is designed and processed for the basic purpose of revenue collection (and not SOI data), we need to consider the quality limitations imposed on SOI data quality by deficiencies present in pre-SOI functions. The keying aspect has already been mentioned. Other areas include non-response (taxpayers who do not file returns or omit data from or report incorrect data on tax returns) and IRS processing functions other than keying, such as controlling of documents and validation of data [12-14].

Application of Quality Control and Management Techniques.--Recognizing that continuous advances are being made in statistical quality control techniques which result in greater effectiveness and efficiency, we plan, through training, to make managers, supervisors and technicians aware of these and encourage their application in SOI processing. (See for example [15-23]). We are already making extensive use of the computer to improve the quality and

efficiency of SOI processing; e.g., computer review and data imputation, as described earlier. The full computer capabilities involving quality data analysis, evaluation, and display have not been realized (See [24].) We hope to remedy this deficiency through quality assurance software training and application in various phases of SOI processing. At the same time, we will take another look at our currently-used computer review system to identify its limitations both as to items it is unable to cover and situations where it is actually introducing errors.

#### ACKNOWLEDGMENTS

The authors would like to thank Richard Thompson for providing much of the data and performing extensive analyses of the 1982 and 1981 1040 SOI Program QM samples. Appreciation is also extended to Norman Waits, Ray Wolfe, Charlie Hicks, and Marshall Epstein of the Returns Analysis Section of the Division's Individual Statistics Branch for their assistance in subject matter and to Ray Shadid of the Branch's Operations Section for his assistance in the sampling area. Special thanks go to Denise A. Reeder for typing the manuscript.

#### NOTES AND REFERENCES

- [1] "IRS Can Do More To Identify Tax Return Processing Problems and Reduce Processing Costs," Report GGD-83-8, General Accounting Office, Washington, DC, October 14, 1982.
- [2] Durkin, Thomas M. and Schwartz, Otto, "The SOI Quality Control Program," American Statistical Association Proceedings, Section on Survey Research Methods, 1981.
- [3] Report of the President's Commission on Federal Statistics, Volumes I and II, Washington, DC, 1971.
- [4] Scheuren, Fritz, Schwartz, Otto and Kilss, Beth, "Statistics from Individual Income Tax Returns: Quality Issues and Budget Cut Impact," Review of Public Data Use, 12:55-67, 1984.
- [5] Deming, W. Edwards, Quality, Productivity, and Competitive Position, Massachusetts Institute of Technology, Cambridge, MA, 1982.
- [6] Data from the 1982 Sole Proprietorship SOI study provides evidence to conclude that the increased production rates for revenue processing at the service centers resulted in a substantial increase in the number of business returns assigned a Principal Industry Activity code for "not allocable." This problem has a serious impact on the 1982 census which makes use of IRS industry data. See also [7].
- [7] Statistical Uses of Administrative Records: Recent Research and Present Prospects, Volume I, U.S. Department of the Treasury, Washington, DC, 1984.
- [8] The Preproduction sample is a batch of returns that include most of the various situations that are expected to be encountered during regular SOI processing. The same batch of returns is processed at each of the service centers (following training) and sent to the SOI Division for review. Results are fed back to the service centers for remedial action.
- [9] Sailer, Peter, Hicks, Charles, Watson, Dave and Trevors, Dan, "Results of Coverage and Processing Changes to the 1980 Individual Income Tax Program," American Statistical Association Proceedings, Section on Survey Research Methods, 1982.
- [10] "Hospital Bills Audited," Equifax Services, Inc., Atlanta, GA, April 30, 1984.
- [11] An error in transcription generated another error at error resolution.
- [12] Some of these conditions, as they apply to income not reported to the IRS, are sometimes referred to as the "Underground Economy." See [13].
- [13] Income Tax Compliance Research, Estimates for 1973-1981, Research Division, Internal Revenue Service, U.S. Department of the Treasury, Washington, DC, July 1983.
- [14] Usually about 0.5 percent of the returns are unavailable for statistical processing.
- [15] Hunter, J.S., "The Technology of Quality: Theory SIGMA," Princeton University, Princeton, NJ.
- [16] Hahn, Gerald J., "Designing Experiments in the Real World —What They Don't Tell You in the Classroom," Presented at the Ellis R. Ott Conference on Quality Management and Applied Statistics in Industry, New Brunswick, NJ, November 1-2, 1983.
- [17] Box, G.E.P., Hunter, W.G. and Hunter, J.S., Statistics for Experimenters, John Wiley, New York, NY, 1978.
- [18] Box G.E.P. and Draper, N.R., Evolutionary Operation, John Wiley, New York, NY, 1969.
- [19] Ott, Ellis R., Process Quality Control Trouble-Shooting and Interpretation of Data, McGraw Hill, Inc., 1975.
- [20] Lucas, James M. and Crosier, Ronald B., "Robust Cusum: A Robustness Study for Cusum Quality Control Schemes," E.I. Du Pont de Nemours and Company, Wilmington, DE.
- [21] Lucas, James M. and Crosier, Ronald B., "Fast Initial Response for CUSUM Quality Control Schemes: Give Your CUSUM a Headstart," Technometrics, Volume 24, Number 3, August 1982.
- [22] Lucas, James M., "Combined Shewhart-Cusum Quality Control Schemes," Journal of Quality Technology, 14, 15-59, 1982.
- [23] Lucas, James M., "Counted Data CUSUMS," E.I. du Pont de Nemours and Company, Wilmington, DE.
- [24] Naus, Joseph I., Data Quality Control and Editing, Marcel Dekker, Inc., New York, NY, 1975.