

RELEASED 10/28/76

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON, D.C. 20548



LM100252



00252

B-137762
Account
by the

OCT 20 1976

The Honorable Charles A. Vanik
Chairman, Subcommittee on Oversight
Committee on Ways and Means
House of Representatives

Dear Mr. Chairman:

Your letter of October 2, 1975, requested that we study the timeliness of Statistics of Income (SOI) data. We have studied the timing problems involved in producing the two most important reports in the series (those for individual and corporation income tax returns) and have prepared answers to your specific questions. A summary of our study, a short background description of the SOI process, and the answers to your questions are presented below.

SUMMARY AND CONCLUSIONS

In brief, SOI publication schedules have considerable built-in lags. Even these generous schedules often are not met. We concluded that, although some delay between the income year and the filing and processing of the returns is unavoidable, some of the lag is due to such inefficient and questionable procedures as the manual typing of all tables for printing, searching repeatedly for errors, and including nonessential features in the reports (such as color printing and some textual material). Perhaps the most important reason for the lag lies in the practice of giving lower priority to a timely product than to such other factors as the quality of the data, the writing and reviewing of the text, and the appearance of the report.

The Internal Revenue Service (IRS) is aware of many of these problems and has made some progress in improving efficiency within the past year. Possibilities for further improvements exist. For example, the schedules could be tightened considerably by adopting some form of computer composition for printing the SOI tables, as most other statistical reporting agencies have done. The reports could be released sooner by eliminating special features that tend to delay publication. Other possibilities which could, with sufficient study and attention, contribute to a more timely product include revision of the time frame of the samples and redefining the quality standards. Finally, the need to

publish SOI on a firm schedule should be recognized as an important goal by the IRS management, and priorities in processing different types of SOI data should be reordered with this in mind.

Another approach would be to consider other IRS data which could be made available earlier than SOI. Requirements for early data on tax payments, for example, might be met by substituting tax administration data for similar items from the SOI sample.

We believe that, in the long run, these improvements will not involve any substantial additional costs. Computer composition of the tables, for example, would involve considerable startup costs but should later prove less expensive than manually typing the tables. We do not recommend the adoption of more costly alternatives until some of the possibilities suggested here have been tried.

BACKGROUND

The Internal Revenue Service accounts for the tax liabilities of taxpayers by making computer tape records (called "transaction tapes") of each tax return as it is filed and posting these to "master files" containing the permanent records of all taxpaying entities. The records originate in the IRS service centers where the returns are filed. For individuals, the initial records contain most of the information found on the income tax return; for corporations, only a few items are included on the transaction tapes. After mathematically verifying the tax computation, scoring for audit potential and other such administrative operations, and posting to the master files, these transaction tapes are erased. It is from these tapes, before posting to the master files, that the statistical samples are selected.

When a tape record is designated for the SOI samples, the return is located and reprocessed; for corporations, all return items needed for the statistical programs are extracted; for individuals, statistical items already on tape (for master file processing) are verified and additional items abstracted. New tape files are created from the sample returns and are processed at the Detroit Data Center. They are tested for consistency of internal relationships, and any inconsistencies are corrected. For each sampling stratum, population counts from the transaction tapes and sample counts from the SOI tapes are used to compute weights; each SOI record is weighted to represent its share of the population from which it was drawn. Analytical tables are produced

and reviewed for internal consistency, trends, and comparability with other data sources. The final output comprises "tax model" tape files and computer printouts of tables. For published reports, the tables are typed (by outside contractors), explanatory text is written and reviewed (by IRS), and the report is printed (by the Government Printing Office or a contractor). Planning, sample design, and all work other than returns processing are under the direct or indirect control of the Statistics Division, IRS.

Table I shows the times scheduled by the Statistics Division for the processing steps for individual and corporation reports for a typical year. The schedule is what the Statistics Division hopes to accomplish; it does not necessarily correspond to the achieved timing for any year. Some timing problems are involved in the schedule itself; others arise in deviations from the planned schedule. Both types of problems are discussed in the following sections.

WHAT IS THE PRESENT STATUS OF SOI REPORTS?

Table II lists the SOI reports in process on June 30, 1976, and shows the date IRS estimates they will be published. Seven of the reports were originally scheduled to have been published before June 1976; two were published in June and the rest were rescheduled to later dates. The reports which have not been rescheduled are not necessarily on time; in fact, two reports due to have been published in July had not been released by early August and at least two others scheduled for this year (the 1973 reports for corporation and business income tax returns) have already encountered difficulties and will almost certainly be rescheduled.

Five SOI reports were published from January through June 1976. The 1971 Corporation Report was published in March, 18 months later than originally scheduled. The 1972 Personal Wealth Study was published in March also, 14 months late. The 1972 Business Returns Report, one of those published in June, was 10 months behind schedule. The 1973 Preliminary Report for Corporations, published in March, and the 1974 Preliminary Report for Individuals, published in June, were 6 and 4 months late, respectively.

WHAT IS THE RECENT HISTORY OF
THE TIMELAG BETWEEN DATA
COLLECTION AND PRINTING OF THE DATA

Table III gives the Statistics Division's calculation of the lag between the filing of the returns and the

publication of the SOI reports for recent years. (By way of comparison, table I shows a lag of 19 months for the individual report and 27 months for the corporation report.) The primary use of table III is to allow comparisons between years; the actual value of the various timelags is not important. This is because the "filing date" is not a particularly meaningful milestone in the production of SOI. While the individual returns for an SOI year are due to be filed by April 15 of the following year, the time to file can be extended to October 15, and delinquencies and foreign residences can extend it even further. Naturally, there is some lag between the time the return is filed and the time it is available for statistical processing. For corporations, both of these problems are greater; a return to be included in the 1975 SOI may be legally filed as late as March 15, 1977, and the lag due to the tax accounting processing of a large corporation return can be serious.

WHAT ARE THE REASONS FOR THE TIMELAG?

1. The filing periods of the returns covered by the reports, as mentioned above, are an obvious, and to some extent unavoidable, source of delay. In addition, however, the Statistics Division's sampling frame allows an extra 2-1/2 months for individual returns and 3-1/2 months for corporations.

2. Getting all the returns designated for the sample is a constant problem, especially those for large corporations, whose returns are likely to be in great demand by audit activities and other IRS elements. In fact, the June cutoff for sampling corporation returns does not apply to very large returns, which are frequently not processed for SOI until the following fall.

3. During most of the period covered by table III (1963 through 1973), both individual and corporation returns were processed for SOI without reference to the data already extracted for the master files. This double processing, which has been modified for individual returns beginning with 1974 SOI, certainly increased costs and probably contributed to the time consumed by SOI processing.

4. Though not apparent from the schedule in table I, finding and correcting errors and inconsistencies is a source of considerable delay. It is, in fact, one of the most important reasons for deviations from the Division's established schedule.

5. The writing and review of explanatory text for the reports also causes frequent departures from established schedules.

6. The copy preparation of the tables for publication is a source of delay built into the schedules (and sometimes causing additional delays). The tables are typed, proofed, and corrected manually. (It is one of the last statistical reports of this magnitude in the United States to be prepared on a typewriter.)

7. Printing consumes a significant amount of time, and some of this may be due to the present design or style of the volumes. The reports contain voluminous tables on many subjects (some of only specialized interest), text (also not always of general interest) illustrated with color charts, a lengthy glossary of terms, and many pages of reproductions of the tax forms. All of these features, however desirable for other reasons, add to printing time. The use of color, for example, is said to add 4 weeks to the schedule.

8. A comparison of the timelag allowed by the Division's optimum schedule (table I), 19 months for individuals and 27 months for corporations, with the actual timelags shown in table III indicates that the Statistics Division has not met its own schedule for the individual reports since the report for 1969 and for the corporation reports since at least 1963. Since the lags have persisted for so long without correction by management, we hesitate to conclude that the above reasons fully account for the observed timelags. We feel that a management audit of the Statistics Division and its activities would be required to answer this question completely.

WHAT STEPS CAN BE TAKEN TO REDUCE THE LAG
OR DEVELOP REASONABLY ACCURATE PRELIMINARY
ESTIMATES AT AN EARLIER DATE?

IRS has already taken some steps which should improve the timing of SOI reports. A task force which has been studying statistical processing for several years has initiated a program for making greater use of existing master file data, has suggested several improvements in returns processing at the service centers, and has initiated some change in the Statistics Division's organization. These changes were designed more to increase efficiency and reduce costs than to improve schedules, but they are also steps in the direction of a more timely product. IRS is

also aware of the problem of locating sample returns; in fact, the Internal Revenue Manual gives SOI priority over most uses of returns. We believe, however, that improving SOI schedules will require IRS to enforce this priority (at least for large corporation returns) far more strictly than in the past.

An additional step which IRS has been contemplating for some time, the use of computer composition to produce the published reports, could save a month or more in the schedules as well as reduce costs. Other agencies, notably the Social Security Administration, have had great success with the use of the Government Printing Office computer composition process for statistical tables. Alternatively, the tables could be redesigned to allow photooffset printing of the computer output, as in most Census publications. We are sending a letter to the IRS Commissioner offering suggestions for improvements in SOI processing, and we are recommending that IRS immediately adopt some system which eliminates manual typing of the tables. (As requested by your office, we have not submitted this report to IRS for comment.)

We have several suggestions which appear to offer possibilities for reducing the time lag but which require further study to determine their feasibility and relative cost. Our letter to the Commissioner suggests that the following changes be considered.

1. Use master file data from the transaction tapes as a substitute for some SOI uses. Since master file data is used to establish tax liabilities and to produce refund checks and bills for balances due, the master file tapes should be a better source of some data--e.g., tax payments--than SOI, as well as being available sooner.

2. Redefine the sampling frames to include individual returns filed from November of one year through October of the next and corporation returns (with the proper accounting periods) filed from July of one year through March of the second year following. This would cover virtually all timely filed returns for the appropriate accounting year. The usefulness of this change would depend on the matching advancement of subsequent steps and also on the ability of the data processing system to assure the availability of large corporation returns.

3. Establish quality standards for SOI data within the time constraints imposed by the schedule. This might

require a cost-benefit study of the present error correction procedures, which appear to have as their goal perfection rather than a reasonable level of quality within a reasonable time.

4. Reduce the time spent writing and reviewing text by curtailing the text itself. A survey of users to determine the usefulness of the present text should be considered.

5. Study the design of the reports with a view to reducing printing time. Consider user need for such features as color charts, the reproductions of return forms, and tables and text of less general interest. Many users, including Treasury's Offices of Tax Analysis and Revenue Sharing, the Joint Committee on Internal Revenue Taxation, and the Department of Commerce, already receive copies of tables on specific subjects well in advance of official publication. Other groups, such as the Budget Committees and the Congressional Budget Office, may also have a need for the information before official publication. Consideration should be given to the possibility that a series of small, topical pamphlets, published as soon as possible after data compilation, could replace some of these pre-publication releases and either replace or supplement the present single, large volumes. Again, users should be consulted, and the design of the report tailored to their needs.

Finally, we believe that some of the problems in SOI arise from conflicting goals and priorities. Users make many different demands on IRS' statistical resources, and IRS is not always in the best position to resolve these demands. An explicit statement of desired publication dates and quality levels, which reconciles the demands of the principal users, would make it much easier for both IRS and any future reviewers to evaluate IRS' statistical efforts. A committee of users, sponsored by the Office of Management and Budget, is currently studying this question, but the timing problem is not its primary concern. We believe that the Office of Management and Budget and the Treasury Department should place a high priority on more expeditious publication of SOI for the benefit of congressional and other users.

WHAT IS THE PROBABLE COST OF SUCH ACTIONS?

We have not performed the cost analyses necessary to answer this question because we feel that cost is not a controlling factor in the sort of changes suggested. We have deliberately avoided alternatives which would involve

substantial increases in costs (such as a larger work force), because such alternatives would do little to improve the timing of the reports without the suggested systems and management changes. We can, of course, make some judgements about the probable costs or savings of individual suggestions, though we believe that the alternatives offered, taken together, can be accomplished within the present budget.

The suggestion for computer composition of tables would involve the cost of training (or contracting for) computer programmers, but it would eliminate the present, very expensive system of typing and proofing every figure published in the reports. After the first year, it could be less expensive. The alternative suggestion for avoiding typing, by printing from photographic reproductions of computer printouts, would be cheaper to implement but would require either increased printing costs or curtailment of the amount of data in the reports.

We see the following cost implications of these suggestions.

1. Greater use of master file data. We envision a small amount of programming and analysis costs that would be offset to some extent by dropping selected items from SOI.
2. Redefinition of sampling frame. Depending on how the speedup in processing was implemented, this could involve small additional costs.
3. Redefinition of quality standards. We believe that this would result in substantial savings.
4. Curtailement of text and redesign of reports. Almost any effort which reduced copy preparation and printing time by reducing the size and complexity of the volumes would also reduce costs. However, a major reorientation of the reports, such as the suggested pamphlet series, could prove costly unless implemented with great care.

B-137762

Finally, management's resolving the conflicting priorities and goals of SOI would, in our opinion, result in a generally more efficient use of resources and would save on costs as well as time.

Sincerely yours,


ACTING Comptroller General
of the United States

Enclosures - 3

TABLE I

STEPS INVOLVED IN PRODUCING INDIVIDUAL AND CORPORATION
STATISTICS OF INCOME FOR 1974

<u>Function</u>	<u>Timespan</u>	
	<u>Individuals</u>	<u>Corporations</u>
1. Sample design	Aug. to Oct. 1973	Aug. to Oct. 1973
2. Program formulation	Jan. 1974 to May 1975	Jan. 1974 to Sept. 1975
3. Returns filed (sampling frame)	Jan. to Dec. 1975	Jul. 1974 to Jun. 1976
4. Sample returns selected and processed	Jan. 1975 to Feb. 1976	Jul. 1974 to Oct. 1976
5. Testing and correcting return records	Oct. 1975 to Feb. 1976	Dec. 1975 to Oct. 1976
6. Early (unpublished) data produced and delivered	Nov. 1975	Mar. to May 1976
7. Preliminary report produced and published	Nov. 1975 to Mar. 1976	May to Sept. 1976
8. Complete report		
a. Tables run	Feb. to Mar. 1976	Nov. 1976 to Jan. 1977
b. Tables reviewed and typed	Mar. to Aug. 1976	Dec. 1976 to Aug. 1977
c. Text written and revised	May to Aug. 1976	Jun. to Aug. 1976
d. Printing	Aug. to Nov. 1976	Sept. to Dec. 1977

TABLE II
STATUS OF STATISTICS OF INCOME PUBLICATIONS

June 30, 1976

<u>Publication</u>	<u>Starting date (note a)</u>	<u>Current estimated completion date</u>	<u>Originally scheduled completion date</u>
1968; 1972 International income taxes (note b)	8/67; 8/71	12/76	6/72; 6/76
1972 Small area data for individual returns (note c)	d/7/73	8/76	10/74
1972 Business income tax returns	8/71	e/6/76	8/75
1972 Corporation income tax returns	8/71	9/76	11/75
1973 Individual income tax returns	8/72	7/76	9/75
1973 Business income tax returns	8/72	8/76	Unchanged
1973 Sales of capital assets, individuals	8/72	1/77	4/76
1973 Corporation income tax returns	8/72	12/76	11/76
1974 Preliminary individual income tax returns	8/73	e/6/76	2/76
1974 Preliminary business income tax returns	8/73	7/76	Unchanged
1974 Small area data from individual returns	d/5/74	9/76	Unchanged
1974 Preliminary corporation income tax returns	8/73	9/76	Unchanged
1974 Individual income tax returns	8/73	10/76	Unchanged
1974 Fiduciary income tax returns	8/73	12/76	Unchanged
1974 Private foundation returns	8/73	3/77	Unchanged
1974 Business income tax returns	8/73	5/77	f/12/77
1974 Corporation income tax returns	8/73	12/77	Unchanged
1974 International income and taxes	8/73	3/79	Unchanged
1975 Preliminary individual income tax returns	8/74	2/77	Unchanged
1975 Preliminary business income tax returns	8/74	7/77	Unchanged
1975 Preliminary corporation income tax returns	8/74	9/77	Unchanged
1975 Individual income tax returns	8/74	11/77	Unchanged
1975 Exempt organizations returns	d/2/75	3/78	Unchanged
1975 Business income tax returns	8/74	5/78	f/8/78
1975 Corporation income tax returns	8/74	12/78	Unchanged
1976 Individual income tax returns	8/75	11/78	Unchanged
1976 Estate tax returns	8/75	1/79	Unchanged
1976 Deferred compensation plan returns	8/75	3/79	Unchanged
1976 Business income tax returns	8/75	5/79	Unchanged
1976 Corporation income tax returns	8/75	12/79	Unchanged

a/ Date of sample design, except as noted.

b/ This is a combination of two separately established projects.

c/ This project is from IMF data and is being programmed for computer composition of the tables.

d/ Date of form establishing project. Work probably began at least 2 months earlier.

e/ Published in June 1976.

f/ Date revised due to Statistics Division decision not to include corporation returns.

TABLE III

NUMBER OF MONTHS ELAPSED BETWEEN FILING DATE AND
PUBLICATION DATE OF STATISTICS OF INCOME REPORTS
FOR INDIVIDUALS AND CORPORATIONS

SOI income year	Individuals		Corporations	
	<u>Preliminary report</u>	<u>Complete report</u>	<u>Preliminary report</u>	<u>Complete report</u>
1963	15	21	19	42
1964	14	21	22	41
1965	10	15	14	30
1966	9	15	12	31
1967	9	18	19	32
1968	13	18	14	30
1969	10	18	15	29
1970	10	23	15	31
1971	11	20	14	a/41
1972	11	21	14	a/33
1973	16	a/23	a/16	a/26
1974	a/10	a/19	a/12	a/26

a/ Estimated.

- Note: 1. For individuals, the SOI income year is the calendar year ending December. Months elapsed are measured from the normal filing date, April 15, of the year following.
2. For corporations, the SOI income year covers accounting periods ending during the span of months, July of one calendar year through June of the following calendar year. Months elapsed are measured from September 15, the normal filing date for corporations with the last included accounting period (ending in June). (This does not take account of returns filed with 3 to 6 months extensions of time, which may be considered timely when filed as late as March 15, following the September 15 filing date.)

Source: Statistics Division, IRS