RESULTS OF COVERAGE AND PROCESSING CHANGES TO THE 1980 INDIVIDUAL STATISTICS OF INCOME PROGRAM

Peter Sailer, Charles Hicks, Dave Watson, and Dan Trevors
Internal Revenue Service

This paper reports on alternative methods employed to close out the statistical processing of the Tax Year 1980 sample of individual income tax returns (Forms 1040/1040A) used for the Statistics of Income (SOI) program. Particular emphasis will be given to the impact of these alternatives on the early "Advance Data" tabulations that are based on an early cutoff of sample receipts for a given tax year. These early tabulations are required annually, for budgetary and tax policy reasons, by the Treasury Department's Office of Tax Analysis and the Congressional Joint Committee on Taxation and are needed no later than the end of November of the year in which the returns are filed.

Organizationally this paper is divided into three sections: (1) an overview of the evolution of SOI individual income tax return statistical processing system, (2) analysis of results of the early cut-off for 1980 Advance Data production processing, and (3) a discussion of subsequent enhancements and research activities of the Statistics of Income Division for the individual income tax return statistical processing system as well as a few suggestions for areas of possible future improvements.

1. RECENT HISTORY OF SOI PROCESSING

Data processing needs of the Statistics of Income program are accomplished as a byproduct of the IRS responsibility to process the income tax returns filed for tax administration purposes. Accordingly, the processing of a statistical sample from these returns is at best a secondary concern to the Service. To the extent possible, statistical needs are incorporated into the mainstream of revenue processing procedures to minimize disruption of the administrative processing of tax returns and to reap the greatest possible benefits from data already entered into the IRS master file system for administrative purposes.

Currently, the individual income tax return sample is identified from "transaction tapes" prepared by the ten IRS service centers for the Individual Master File (IMF). Prior to Tax Year 1974, all of the data used in the Form 1040 SOI program were manually abstracted or edited onto hardcopy edit sheets from the tax returns as a separate off-line statistical processing operation. For Tax Year 1974, the first attempt was made to use IMF data for SOI. For that year, a limited number of codes and amounts were computer printed onto edit sheets for returns in the sample using the transaction files. These data were subsequently reviewed by statistical clerks for statistical acceptability. At the same time, additional data were needed for statistical

purposes, but were not available from the transaction files. Such data were abstracted from the tax return edited as necessary. As confidence and experience were gained in the usability of the IMF data, more SOI data came to be based on this source. During this evolution, enhancements were added to this new system, such as computer checking the validity of the statistical data while the returns were still on hand. Unfortunately, as enhancements were added, the ability to meet interim processing deadlines was strained. In fact, meeting some of these deadlines became impossible because of the processing lags these modifications created. Therefore, for Tax Year 1980 a major revision was made to one of the most critical dates.

This change accelerated the sample cut-off date by two weeks, effectively moving up the ending date for inclusion of sampled returns in the Advance Data report from mid-October to the end of September. In terms of the number of returns involved, accelerating the cut-off date by two weeks apparently excluded 2,200 return records from the sample of 160,133 returns. As it turned out though, the number of returns excluded from the sample because of this early cut-off, as compared to those under the previous years' early cutoffs, were approximately the same. Traditionally, any designated sample return records for which revenue processing was incomplete (at that given point of time) were omitted from the Advance Data file, thus resulting in Advance Data tabulations that were actually based on returns designated for the sample by mid-September or earlier.

The new effort, for Tax Year 1980, was designed to assure that all returns designated through mid-September were processed and shipped on tape for statistical processing (at the IRS Data Center, in Detroit) within a time frame that was approximately two weeks earlier than previous years. Most of the "final" returns shipped to the Data Center at cutoff time for inclusion in Advance Data included only transaction file data. Time did not permit any perfecting of these data, or the manual editing of data not available from the transaction file. Because of the missing data, these records would not pass the computerized validity tests used to further process the records into a form from which tables could be produced. Therefore, portions of the missing data had to be imputed and these imputations were limited to those necessary to enable the record to pass all of the tests. The imputed amounts were determined, if possible, from data present on the record. Otherwise, amounts were estimated on a proportional basis, using data available from returns for which the manual editing had been completed. For example, net

capital gain or loss was the amount on the transaction file and was therefore the only capital gain amount available to complete processing. However, details on long- and short-term capital gains, which normally would have been manually edited from the return, were needed. If the alternative minimum tax had been used by a taxpayer, this tax figure was available from the transaction files, and since the long-term capital gain excluded from adjusted gross income is one of the items used to compute alternative minimum tax, the gain amount could be determined by working backwards from the tax. If alternative minimum tax was not present, or the entire gain amount could not be determined using this method, long-term or short-term gain was determined based on the proportion of each on similar returns for which these amounts had been manually edited.

For the later SOI complete report, a more detailed and sophisticated set of imputation factors was developed to cover the manually-edited items for these returns. Again, characteristics present on similar returns were taken into account in making the imputations. If these proved inadequate, then distributions were made on a proportional basis using as a guide the 1980 "Advance Data," but only for fully processed returns.

The following section describes the results of the early cutoffs for Tax Year 1980. It should be noted that these procedures were not undertaken simply on faith. Rather, a series of tables was produced, using the 1978 Statistics of Income File, in which the dropping of late-filed returns and the weighting of earlier ones was simulated. In a

paper written about this simulation [1], James Dumais and Ray Shadid concluded that a cycle 36 (mid-September) cutoff should be adequate for producing basic income and tax estimates, and that an earlier cutoff for the Complete Report was also feasible.

2. ANALYSIS

This section presents an analysis of the effects on Advance Data and on SOI of advancing the closeout date and including imputations and "unedited" transaction file data in the statistics for 1980. The tabulations discussed are shown at the end of this paper.

It is important to note that these two strategies for expediting the data actually involved only a very small proportion of the total sample. When the transaction file data are used in combination with imputations, the manual editing step can be bypassed and the return record processed immediately through the service center directly to the IRS Data Center. If the return comes in very late, it may not be included at all in the file used to prepare the Complete Repoort. However, its absence will be offset by weighting the existing returns already on file. As is shown in figure A, only 0.9 percent of the final estimate for adjusted gross income was based on returns processed in this manner; only 0.2 percent of the final estimate was derived by assigning higher weights to sampled returns to compensate for returns not yet sampled. Not surprisingly, returns "forced" through the system were concentrated in the highest income and in the deficit classes.

For the Advance Data closeout, the

Figure A.—Complete Report: Percent of Total Estimate
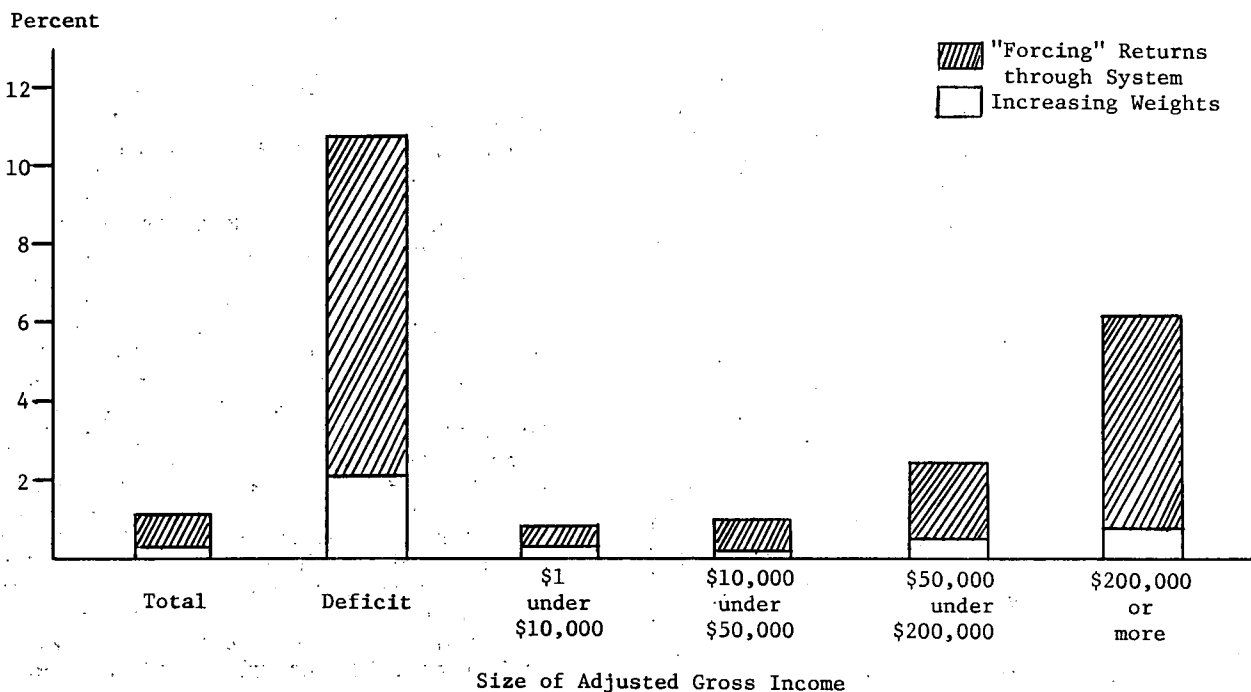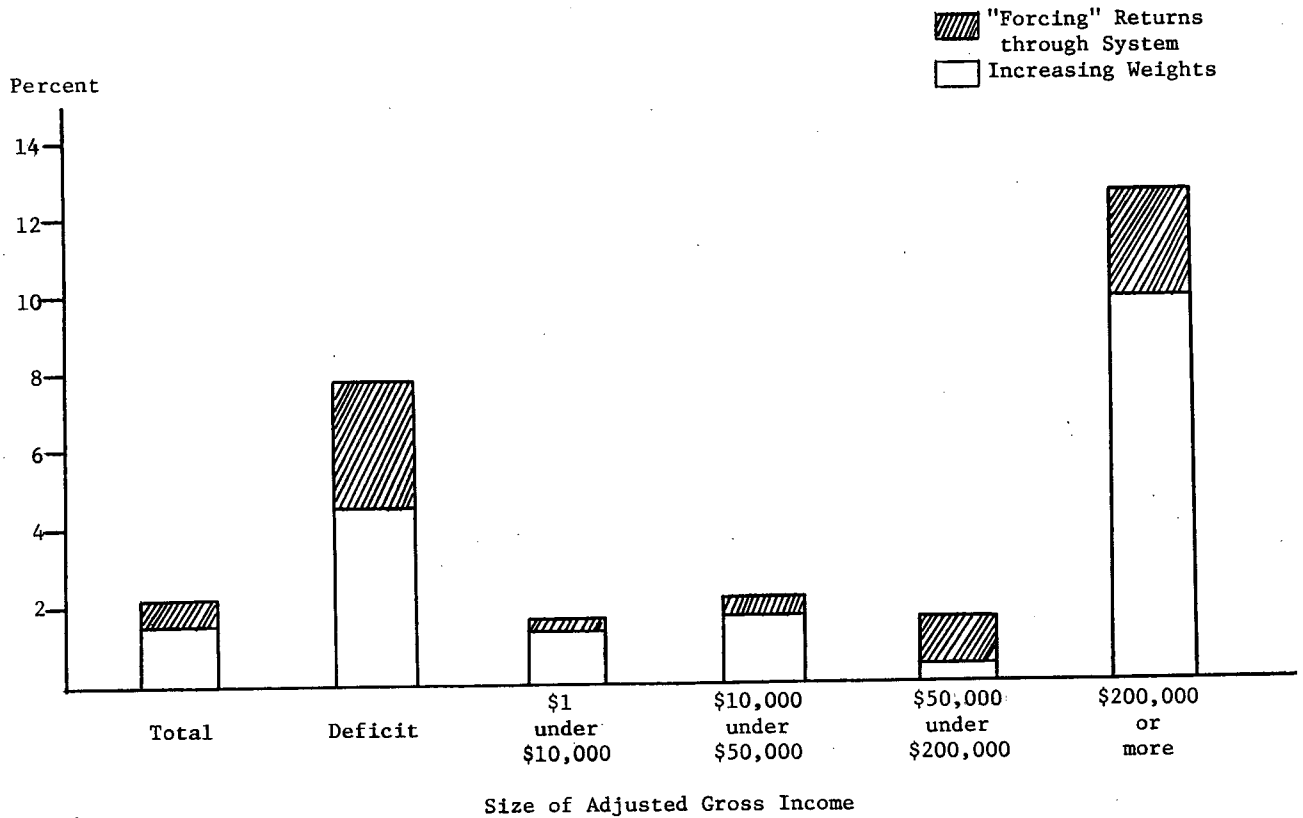of Adjusted Gross Income by Estimation Method

Figure B.--Advance Data: Percent of Total Estimate
of Adjusted Gross Income by Estimation Method

proportion of returns forced through the system was slightly lower than that for the later SOI Report--0.6 percent vs. 0.9 percent for the SOI Report. Obviously, as can also be seen from Figure B, the portion of the Advance Data estimates derived by weighting early returns to replace later returns was much higher--1.5 percent overall (0.2 percent for the SOI Report); 9.8 percent for high-income returns (0.6 percent for the SOI Report).

The strategy for expediting publication of data for 1980 consisted simply of moving up the closeout date by two weeks for Advance Data (to mid-September) and by four weeks for the later SOI Report (to the end of November). Since the number of returns processed for a year are so close to completion by mid-September, let alone by late November, it would not appear that cutting off two weeks earlier would have a significant impact on the results. In fact, all of the items in the SOI Complete Report (prepared from the file with the late November closeout) were within a fraction of a percent of the data from the final file, as is shown in Table 1. (The final file was one created after all returns selected for the sample during Calendar Year 1981 had been processed through the system.) As would be expected (given the overlap between the two samples), all of the differences shown in Table 1 are much less than the expected sampling variability at the one standard deviation level of significance.

Table 2 shows a comparison between all of the items customarily shown in Advance Data Reports with the comparable items from the corresponding Complete Reports, for Tax Years 1976 through 1980. For the two most basic data items -- adjusted gross income and total tax liability -- the average differences between the early and Complete Report estimates were 0.13 and 0.23 percent, respectively. For 1980, the differences were 0.5 percent for adjusted gross income, and 0.8 percent for total tax liability, a moderate increase over the previous years, but still, it would appear, perfectly acceptable as early estimates. (It should be noted that the increased differences are due only in part to the earlier cutoff for returns processing. An unexpected surge in the number of returns with very high income filed during December of 1981 also led us to understate the weights for the top income class in producing the 1980 Advance Data estimates).

However, while the results were reasonably encouraging for the basic items, certain rarer items presented a real problem. Table 2 also compares Preliminary and Complete Report data for two items that are relatively rare, that are somewhat complicated for the taxpayer to compute, and that have been subject to frequent tax law changes in recent years: the minimum tax and the alternative minimum tax. Quality of the data for these items for the years with an October 1 Advance Data cutoff was relatively

poor. The mid-September cut-off for 1980 appears to have made matters worse.

Obviously, even though a very small percentage of returns is filed after mid-September, the characteristics of these returns differ sigificantly from those of the other returns, so that weighting the other returns to compensate for the returns filed later in the year is not the complete answer.

The differences discussed so far are attributable to the fact that late returns tend to be different from early returns. One reason for the difference is that returns filed later in the year tend to be more complex than those filed earlier. This can be seen clearly from Figure C, which shows the level of complexity of returns by the week in which they were processed. The measure used to indicate complexity is the number of schedules attached to the basic Form 1040.

As can be seen from Figure C, the median value of this measure of complexity remains at less than one (schedule) through the end of March (week 12 of the Processing Year); then rises steadily through mid-July (week 27). Between mid-July and mid-September (weeks 27 through 36), it remains fairly constant at between five and six (schedules), and then rises once more to between six and seven. The median value for the year as a whole is 3.5 (schedules).

The first recommendation for future improvement is to return to an October 1 cutoff, and make up the two weeks in other ways -- specifically, by reducing the amount of time it takes to edit returns and by printing out from the transaction files computer-generated codes that tell the editors precisely what other information to look for on the returns, based on the data available from the transaction files, and which data items need to be
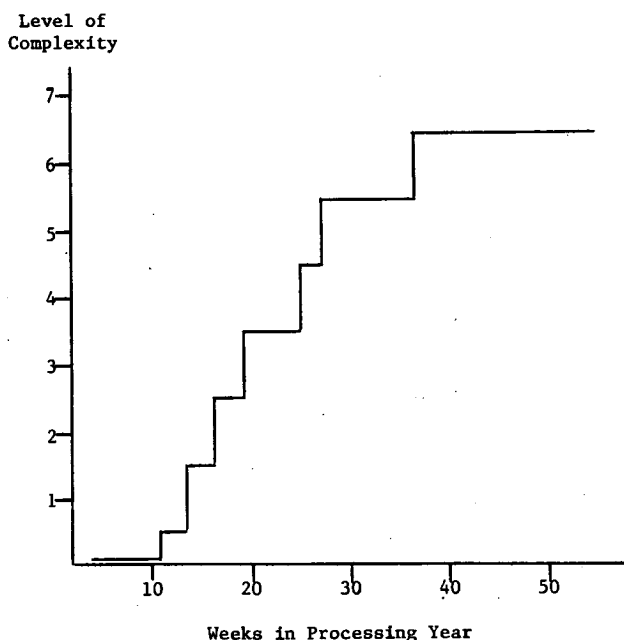
separately abstracted or edited from the return. Thus, for instance, a code based on the data available from the transaction files could advise the editor to go to the investment credit schedule to obtain additional data, based on the presence of the investment credit on the transaction file. Such a system has been instituted for tax year 1981, as is further explained in the next section.

The second recommendation is not to weight the very earliest (and simplest) sample returns to compensate for the later, complicated ones. Rather, we should increase the weights primarily on the returns that come in after June 1, which is the time by which most timely-filed returns have been processed. The trade-off here is that, while reducing the bias that is built in by weighting early returns to make up for late ones, we would also increase the sampling variability of the estimates by applying unusually large weights to a small group of late returns. Nonetheless, when looking at an estimate like alternative minimum tax which is off by 28 percent, even though expected sampling variability is only 3 percent, the trade-off appears acceptable.

The second strategy for expediting processing was the so-called "forced" processing of data from the IMF directly into the statistical file. What was sent to the IRS Data Center in these cases was a record that had only transaction file data. Since they comprised a very small proportion of the total sample, the data for these returns had little impact on the quality of the totals presented. However, there are a number of items needed for SOI not available from the transaction files and, for the "forced" returns, these items had to be imputed. Later, a small subsample of these returns was retrieved and fully processed in order to evaluate the quality of the imputations. These differences are summarized below.

Figure D shows that, for items where the IMF transaction data offered a number of clues which improved the imputations, the imputations were quite valid. For instance, in the case of net long-term capital gains, we had the advantage that the missing item had to balance not only to net capital gain in adjusted gross income, but also to the alternative minimum tax. For many returns, there was in effect only one plausible imputation and the method of imputation used proved quite accurate. On the other hand, the imputed credit card interest deduction was based on the proportion of the total interest deduction that came from this source on returns processed earlier in the year. This proportion turned out to be invalid because the returns processed later in the year tended to have much larger interest deductions, of which credit card interest was a much smaller proportion. If an imputation for credit card interest is needed for future SOI programs (it is not part of the 1981 SOI program), the percentage will need to be varied by income class and also possibly by processing week [2]. It is important to note that, even though the imputations for such items as credit card interest and net short-term capital loss were not very accurate, because the imputed

Figure C.--Median Level of
Complexity by Processing Week

Level of
Complexity

Weeks in Processing Year

Figure D. Imputed Data as a Percent
of Fully Processed Data

Item	Imputed as a Percent of Final Estimate 1/	Potential Distortion of Final Estimate 2/
General sales taxes deducted...............	83.6	-0.2
Personal property taxes deducted...............	93.1	-0.1
Credit card interest deducted...............	207.0	0.3
Union dues deducted.....	133.2	0.2
Net short-term capital gain...................	99.6	-0.0
Net short-term capital loss...................	27.6	-4.7
Net long-term capital gain...................	99.3	-0.0
Net long-term capital loss...................	98.4	-0.0
Personal service gross income................	97.2	-0.1
Deductions from personal service gross income...	157.5	0.7
Personal service net income................	96.8	-0.1

1/ Obtained from 103 returns for which data
were first imputed, then abstracted and
keyed under regular procedures.

2/ Assuming the same processing (non-
sampling) error for all the remaining
unedited returns as for the sample of 103,
this is the percentage by which the
published estimate could be off.

amounts represented such a small percentage of
the final estimates, the affect on the quality
of the final estimates was quite small.

In addition to developing improved
imputations for future years, another change
instituted which should improve the quality of
any future "forced" return estimates is the
incorporation of even more data items from the
transaction files into the SOI files. Past
constraints on the number of transaction file
items that could be incorporated into the SOI
system have now been eliminated by other
changes made to the SOI processing system.

In conclusion, we are reasonably satisfied
with the results of the early closeout and
forced processing strategies instituted for the
1980 SOI program. Both strategies were
applicable to only a small percentage of the
total returns and, in terms of cost-benefit,
the gain in timeliness of the SOI data was
considerable.

3. FUTURE RESEARCH EFFORTS

Evolution of the 1040 Statistics of Income
(SOI) system, with its expanded use of

transaction file data, diminishes the need for
a manually prepared edit sheet for every sample
return selected. In fact, unpublished analyses
conducted within the Division indicate that
most Form 1040A data could be accepted at "face
value" from the transaction files without the
need for visual inspection at the service
centers. For Tax Year 1981 (Filing Year 1982),
the structured fixed-design edit sheet has now
been eliminated. Because of the physical
limitations of printing transaction file data
on fixed format edit sheets, only 32 percent of
the data items needed for SOI could previously
be obtained from the IMF. Elimination of the
fixed design edit sheet now allows us to obtain
63 percent of the total SOI items from the
IMF. Instead of the edit sheet, data are read
out onto ordinary computer printout paper, but
only if computer tests or checks determine that
there are data inconsistencies or indicate that
additional data need to be obtained from the
return by the statistical editors. Return
records that pass the tests and checks and
consequently require no additional statistical
editing are processed directly to the tapes to
be sent to the Data Center. Figures through
June 1982 show that approximately 28 percent of
the 1981 returns selected for SOI have been
processed directly to the Data Center tape
file. Returns processed within this time frame
are the "early filers" and consist typically of
Form 1040A and the simpler Form 1040 returns.
During the last half of the year the more
complex and prior-year Form 1040 returns will
be processed by the Service and then selected
for the SOI sample. Thus, a decrease in the
percentage of returns (records) processed
directly to the Data Center file can reasonably
be expected. Under the new system, statistical
editors can now focus their attention directly
on those returns requiring review and only to
the specific area or areas within the return
record in need of scrutiny. In addition, the
amount of time spent on batching and con-
trolling has been reduced significantly.

Simplified systems and record design for
source data capture permit simplified tran-
scription techniques to be used. IRS employs a
unique direct data entry system (DDES) designed
for the administrative processing of tax returns
that possess generalized parameter driven (GPP)
transcription capabilities for any off-line
processing. GPP, while effective, is not
necessarily efficient. It is however, the
medium that must be used to meet SOI transcrip-
tion needs. Conversion to the equivalent of an
unstructured edit sheet and resultant simplifi-
cation of the processing system has resulted in
an approximate 600 percent increase in the
transcription rate. Nominal resources are also
saved in the areas of paper, printing, and
computer time. Eliminated then is the
structured edit sheet whose function had
evolved from that of an essential source data
capture document to one of an intermediate (or
lesser) role of data display.

For Tax Year 1981, sample receipts for
both Advance Data and SOI report processing
will be cut off early. Unlike Tax Year 1980,
the Advance Data cutoff has been automated to
the extent that data processed through mid-

October, as opposed to early October, can be included without compromising the delivery dates. Plans are to cut off the sample for the SOI Report processing at about the beginning of December. Sample designation will continue, however, through the end of Calendar Year 1982. Any returns thus excluded, but which have data characteristics whose absence from the sample could bias the results, can be introduced into the file during the processing at Detroit. These returns include, but are not restricted to, high income nontaxables, large adjusted gross income or deficit returns, or returns with a large amount for any specified data item.

One further change introduced for the 1981 program is the combined processing, testing, and correction of individual (1040) data and sole proprietorship (Schedules C and F) data [3]. As a result, the previous practice of splitting off these two files, controlling them separately, and then recombining them will no longer be necessary.

Areas for future research include possible telecommunication of data between the service centers and the Data Center as opposed to conventional shipping methods now used. Consideration is also being given to alternative methods of handling prior-year returns, currently included in the sample as "stand-ins" for delinquent current-year returns yet to be filed. If such returns could be eliminated from the sample, a good deal of "exception processing" could be avoided. Finally, feasibility studies for modifying the on-line error resolution (currently planned for tax administration purposes) so that it can be applied and adapted to SOI processing is also under way.

The Statistics of Income Division, like most Federal statistical organizations, is increasingly faced with budgetary constraints. Future budgetary constraints may have to be met by greater use of the IMF data, computerized imputation and correction routines, and elimination of possibly superfluous manual functions. With the increasing trend by taxpayers to file as late in the calendar year as is legally possible, the early cut-off of the sample may prove to be only an interim solution. Long-range strategy indicates that streamlining and standardizing of procedures is the direction in which SOI Division must proceed.

ACKNOWLEDGMENTS

NOTES AND REFERENCES

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

[2] For other possible approaches to imputing missing data, see Hinkins, Susan, "Imputation of Missing Items on Corporate Balance Sheets," 1982 American Statistical Association Proceedings, Section on Survey Research Methods.

[3] Wolfe, Raymond, Methodological Changes in the Statistics of Income Sole Proprietorship Programs--Dominant Business Processing, unpublished working paper available from the Statistics of Income Division, Internal Revenue Service.

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

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

[6] Internal Revenue Service, Statistics of Income--1980 Individual Income Tax Returns, Washington, D.C. 1982.

[7] Internal Revenue Service, Statistics of Income Bulletin, Winter 1981-82, Washington, D.C. 1982.

Table 1--Selected Income and Tax Items, Complete Report and Final File,
Statistics of Income-1980 Individual Income Tax Returns

(All figures are estimates based on samples - money amounts are in thousands of dollars)

Item	1980 Complete Report			1980 Final File			Complete Report as % of Final File
	Number of Returns	Amount	Average	Number of Returns	Amount	Average	
Adjusted Gross Income	93,238,823	1,626,554,501	17,445	93,238,724	1,626,488,454	17,444	100.0
Adjusted Gross Deficit	663,646	-12,823,004	-19,322	663,717	-12,914,356	-19,458	99.3
Salaries and Wages	83,802,109	1,349,842,802	16,107	83,802,041	1,349,818,631	16,107	100.0
Business Net Profit	6,305,794	66,995,010	10,624	6,305,781	66,996,228	10,625	100.0
Business Net Loss	2,575,325	11,865,856	4,608	2,575,333	11,883,280	4,614	99.9
Farm Net Profit	1,123,085	9,938,950	8,850	1,123,079	9,938,728	8,850	100.0
Farm Net Loss	1,485,345	11,731,416	7,898	1,485,351	11,749,976	7,911	99.8
Partnership Net Profit Less Loss	3,154,658	9,428,595	2,989	3,154,535	9,429,859	2,989	100.0
Small Business Corp. Net Profit Less Loss	755,868	670,751	887	755,820	670,892	888	100.0
Net Capital Gain	59,794,437	32,356,419	541	59,794,424	32,400,613	542	99.9
Net Capital Loss	1,957,008	3,066,293	1,567	1,956,997	30,662,159	1,567	100.0
Sale of Property Other Than Capital Assets							
Net Gain	669,735	2,149,695	3,210	669,750	2,159,438	3,224	99.5
Net Loss	424,922	2,073,409	4,880	424,935	2,113,197	4,973	98.1
Total Div. Received	14,567,903	40,375,626	2,772	14,567,847	40,337,247	2,769	100.1
Dividends in AGI	10,738,982	38,761,253	3,609	10,738,927	38,722,884	3,606	100.1
Interest Received	49,019,575	102,009,444	2,081	49,019,538	102,024,527	2,081	100.0
Pensions and Anuities in Adjusted Gross Income	7,373,704	43,339,736	5,878	7,373,679	43,339,561	5,878	100.0
Rent Net Income	3,653,996	13,664,691	3,740	3,653,973	13,681,087	3,744	99.9
Rent Net Loss	3,809,821	13,464,274	3,534	3,809,868	13,528,480	3,551	99.5
Royalty Net Income Less Loss	744,315	3,904,984	5,246	744,274	3,899,645	5,240	100.1
Estate and Trust Net Income Less Loss	906,297	4,560,031	5,031	906,257	4,560,856	5,033	100.0
State Income Tax Refund	10,525,604	3,629,571	345	10,525,601	3,630,008	345	100.0
Alimony Received	317,335	1,422,198	4,482	317,334	1,422,191	4,482	100.0
Other Income	3,633,028	6,799,669	1,872	3,633,048	6,803,189	1,873	99.9
Other Loss	389,511	7,378,903	18,944	389,556	7,418,834	19,044	99.5
Total Adjustments	13,148,919	28,614,061	2,176	13,148,871	28,616,177	2,176	100.0
Disability Income Exclusion	297,720	1,034,404	3,474	297,770	1,034,404	3,474	100.0
Payment to IRA	2,564,421	3,430,894	1,338	2,564,412	3,430,894	1,338	100.0
Payment to a Keogh Plan	568,936	2,007,666	3,529	568,922	2,007,551	3,529	100.0
Deductions for Excess Foreign Living Expenses	119,686	1,299,006	10,835	119,660	1,300,020	10,864	99.9
Exemption Amount	93,902,469	227,925,098	2,427	93,902,441	227,569,162	2,423	100.2
Unused Zero Bracket Amount	497,304	915,851	1,842	497,303	915,848	1,842	100.0
Excess Itemized Ded.	28,634,179	127,935,816	4,468	28,634,083	127,938,245	4,468	100.0

Table 1--Selected Income and Tax Items, Complete Report and Final File,
Statistics of Income-1980, Individual Income Tax Returns--Continued
(All figures are estimates based on samples - money amounts are in thousands of dollars)

Item	1980 Complete Report			1980 Final File			Complete Report as % of Final File
	Number of Returns	Amount	Average	Number of Returns	Amount	Average	
Total Medical and Dental Deduction	19,458,318	14,972,082	769	19,458,265	14,972,511	770	100.0
Total Taxes Paid Ded.	28,749,278	69,404,275	2,414	28,749,186	69,402,908	2,414	100.0
Total Interest Paid Ded.	26,676,836	91,187,006	3,418	26,676,768	91,190,844	3,418	100.0
Total Contribution Ded.	26,601,428	25,809,608	970	26,601,344	25,807,617	970	100.0
Casualty and Theft Ded.	2,943,909	3,513,963	976	2,943,896	3,513,218	1,193	100.0
Total Miscellaneous Ded.	24,034,377	13,141,206	547	24,034,289	13,143,155	547	100.0
Total Itemized Ded.	28,950,282	218,028,139	7,531	28,950,185	218,030,253	7,531	100.0
Tax Table Income	93,039,759	1,500,738,938	16,129	93,039,658	1,500,671,963	16,129	100.0
Income Tax Before Cred.	76,135,819	256,294,315	3,366	76,135,723	256,254,609	3,366	100.0
Total Tax Credits	19,674,483	7,215,839	367	19,674,395	7,211,436	367	100.0
Targeted Jobs Credit	119,082	196,357	1,649	119,088	196,388	1,649	100.0
Earned Income Credit used to offset Income Tax Before Credits	3,154,428	451,366	143	3,154,428	451,366	143	100.0
Residential Energy Cred.	4,669,675	562,141	120	4,669,666	562,142	120	100.0
Income Tax After Cred.	73,840,395	249,078,475	3,373	73,840,290	249,043,173	3,373	100.0
Total Tax Preferences	166,496	5,442,345	32,688	165,985	5,427,173	32,679	100.3
Minimum Tax	94,599	412,638	4,362	94,621	413,414	4,369	99.9
Alternative Minimum Tax	122,670	850,326	6,932	122,696	854,261	6,962	99.5
Total Income Tax	73,906,244	250,341,440	3,387	73,906,166	250,310,848	3,387	100.0
Self Employment Tax	7,723,051	5,659,897	733	7,723,021	5,659,861	733	100.0
Earned Income Credit used to offset all other taxes	720,863	164,461	228	720,864	164,461	228	100.0
Total Tax Liability	75,325,013	256,251,076	3,402	75,324,945	256,220,651	3,402	100.0
Total Taxpayments	87,439,159	271,501,122	3,105	87,439,110	271,488,923	3,105	100.0
Income Tax Withheld	81,727,468	228,959,396	2,801	81,727,420	228,952,371	2,801	100.0
1980 Estimated Tax Payments	8,503,963	37,867,010	4,453	8,503,901	37,834,815	4,449	100.0
Earned Income Credit, Refundable Portion	4,996,637	1,370,169	274	4,996,638	1,370,169	274	100.0
Tax Due at Time of Filing	21,755,516	32,843,576	1,510	21,755,463	32,828,329	1,509	100.0
Total Overpayment	69,868,451	49,458,344	708	69,868,458	49,461,323	708	100.0
Overpayment Credited to 1981 Estimated Tax	2,174,098	2,778,272	1,278	2,174,095	2,777,753	1,278	100.0
Overpayment Refunded	68,232,903	46,680,072	684	68,232,921	46,683,569	684	100.0

82

Table 2--Selected Income and Tax Items, Preliminary and Complete Reports,
Individual Statistics of Income, 1976-1980

(All figures are estimates based on samples--money amounts are in thousands of dollars)

Item	1976			1977		
	Preliminary	Complete	Preliminary as Percent of Complete	Preliminary	Complete	Preliminary as Percent of Complete
Total number of returns	84,536,143	84,670,389	99.8	86,493,391	86,634,640	99.8
Adjusted Gross Income	1,053,592,868	1,053,895,687	99.9	1,155,699,080	1,158,492,225	99.8
Salaries & Wages	879,769,742	880,998,631	99.8	965,608,458	969,403,997	99.6
Dividends in AGI	24,451,749	24,461,943	99.9	27,132,902	27,020,483	100.4
Total Adjustments	15,940,772	16,284,275	97.8	19,032,735	19,329,363	98.5
Total Itemized Ded.	133,218,032	133,890,976	99.5	137,294,496	138,519,421	99.1
Income Tax Before Credits	153,550,742	153,534,155	100.0	171,407,238	172,111,669	99.6
Total Tax Credits	12,631,621	12,732,959	99.2	13,319,588	13,637,761	97.7
Income Tax After Credits	140,919,120	140,801,196	100.1	158,087,650	158,473,908	99.8
Minimum Tax	967,232	1,000,273	96.7	1,243,179	1,322,915	94.0
Alternative Min. Tax	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Total Income Tax	141,886,352	141,801,470	100.1	159,330,829	159,796,823	99.7
Total Tax Liability	145,813,508	145,748,802	100.1	163,513,341	164,024,104	99.7
Total Tax Payments	158,466,861	158,634,719	99.9	176,210,469	176,874,876	99.6
Tax Due at Time of Filing	17,500,583	17,483,628	100.1	19,766,962	19,990,260	98.9
Total Overpayments	31,015,531	31,304,595	99.1	33,275,338	33,715,617	98.7

Item	1978			1979		
	Preliminary	Complete	Preliminary as Percent of Complete	Preliminary	Complete	Preliminary as Percent of Complete
Total number of returns	89,889,669	89,771,551	100.1	92,616,213	92,694,302	99.9
Adjusted Gross Income	1,304,188,847	1,302,447,386	100.1	1,463,666,582	1,465,394,530	99.9
Salaries & Wages	1,092,017,073	1,090,291,855	100.2	1,229,353,731	1,229,251,389	100.0
Dividends in AGI	30,169,755	30,206,475	99.9	33,078,347	33,482,508	98.8
Total Adjustments	22,333,986	22,364,088	99.9	24,685,773	24,778,484	99.6
Total Itemized Ded.	163,807,295	164,432,406	99.6	182,910,289	184,168,669	99.3
Income Tax Before Credits	204,162,283	203,803,653	100.2	219,306,130	220,099,516	99.6
Total Tax Credits	16,989,359	17,085,591	99.4	6,556,492	6,780,186	96.7
Income Tax After Credits	187,172,925	186,718,062	100.2	212,749,638	213,319,330	99.7
Minimum Tax	1,404,261	1,514,475	92.7	274,607	309,248	88.8
Alternative Min. Tax	n.a.	n.a.	n.a.	729,850	865,940	84.3
Total Income Tax	188,577,186	188,232,537	100.2	213,754,094	214,494,519	99.7
Total Tax Liability	193,464,593	193,184,849	100.1	219,244,082	220,100,162	99.6
Total Tax Payments	203,340,183	202,829,400	100.3	236,845,190	237,099,515	99.9
Tax Due at Time of Filing	24,729,595	24,969,333	99.0	26,797,803	27,427,573	97.7
Total Overpayments	35,362,293	35,415,451	99.8	45,804,533	45,821,372	100.0

Footnote at end of table

Table 2--Selected Income and Tax Items, Preliminary and Complete Reports,
Individual Statistics of Income, 1976-1980--Cont.
(All figures are estimates based on samples--money amounts are in thousands of dollars)

Item	1980		
	Preliminary	Complete	Preliminary as Percent of Complete
Total number of returns	93,616,278	93,902,469	99.7
Adjusted Gross Income	1,606,265,685	1,613,731,497	99.5
Salaries & Wages	1,345,004,185	1,349,842,802	99.6
Dividends in AGI	37,971,572	38,761,253	98.0
Total Adjustments	28,025,600	28,614,061	97.9
Total Itemized Ded.	214,784,413	218,028,139	98.5
Income Tax Before Credits	254,211,495	256,294,315	99.2
Total Tax Credits	6,745,156	7,215,839	93.5
Income Tax After Credits	247,466,339	249,078,475	99.4
Minimum Tax	323,296	412,638	78.3
Alternative Min. Tax	610,967	850,326	71.9
Total Income Tax	248,400,602	250,341,440	99.2
Total Tax Liability	254,120,823	256,251,076	99.2
Total Tax Payments	269,764,816	271,501,122	99.4
Tax Due at Time of Filing	32,138,659	32,843,576	97.9
Total Overpayments	49,137,184	49,458,344	99.4

n.a. - not applicable