

Individual Income Tax Returns: Selected Characteristics From the 1980 Taxpayer Usage Study

By Paul Grayson*

Since 1970 the Taxpayer Usage Study (TPUS) of the Internal Revenue Service has provided, shortly after the close of each filing period, early indications of how taxpayers are using the current-year Form 1040 and 1040A Individual Income Tax Returns. In view of changes in the tax laws, redesign of the tax forms and other factors, IRS officials have found this early analysis of the characteristics of income tax returns useful for a variety of planning purposes involving processing workloads, forms design, taxpayer reporting behavior, and the like. The study is based on a sample of about 7,500 Forms 1040 and 1040A filed between January 1 and about May 1 each year. The study report is normally released during the summer, providing frequency information on return entries made by taxpayers, and other key characteristics of the returns filed.

THE SIGNATURE OF A PREPARER

Whether more or fewer taxpayers--and what kind--engage the services of return preparers is significant for a variety of reasons. From the IRS standpoint, a preparer return is likely to be more expensive for IRS to process than a self-prepared one, as evidenced by data presented here. From a broader perspective, the incidence of preparer returns can be viewed as a reflection of the complexity of the tax law and the tax return. From the taxpayers' standpoint, their perceptions of the need to employ preparers result in additional costs to them in complying with the tax laws. In addition, one may view the return preparer as a medium whose cooperation IRS can enlist to improve the tax administration system. What, then, do we know about preparer returns and their characteristics?

The Taxpayer Usage Study employs, as an indicator of a "preparer return," the presence of an accepted signature on the line of the return for "Paid Preparer's Use Only--Preparer's signature and date." Other returns are designated "self-preparer." (Included in this latter category are returns with a signature on the line made by a family member or inadvertently by the taxpayer. Also classed as "self-preparer" are returns signed by identifiable unpaid preparers such as IRS personnel.)

In terms of specific forms filed for 1980, 52 percent of Forms 1040 were signed by preparers as opposed to 17 percent of Forms 1040A. The trend in average usage of preparers over recent years is as follows:

Tax year	Preparer returns (Percent)	Tax year	Preparer returns (Percent)
1980.....	38	1976.....	47
1979.....	39	1975.....	46
1978.....	39	1974.....	43
1977.....	42	1973.....	48

The downward trend of the last four years is accentuated by the drop in 1976-1977. Beginning with 1977, the law has called for the signature of only paid preparers. The 1976-1977 shift can undoubtedly be attributed to the change in the requirement: many unpaid preparers, such as friends and relatives, no longer felt obligated to sign the returns they prepared.

Tables 1 and 2 provide comparative "profiles" of preparer and self-preparer returns filed during January through April of 1981, for Tax Year 1980. Eighty-two percent of preparer returns were filed on Form 1040; only 46 percent of self-preparer returns were 1040's. Preparer returns were more likely to have entries for income tax liability, but less likely to have refunds, than self-preparer. Preparer returns were less likely to have the label, envelope and return that IRS provides in the mailing package--a consistent TPUS finding over the course of several years.

Table 2 shows that certain attached forms and schedules are more frequent among preparer returns than among self-preparer, but that this does not hold for all attachments. Schedule A, for example, accompanied about one-half the 1040's filled out by preparers and about the same proportion of self-preparer 1040's. (The names of these schedules may be found at the end of this article.) Schedules C and E represent the more common condition: each was relatively more frequent among preparer than self-preparer Forms 1040.

Table 1.--Percentage Frequency of Selected Items Among Preparer and Self-Preparer Returns, Tax Year 1980

Item	Individual returns		
	All	Preparer	Self-preparer
	(1)	(2)	(3)
Form 1040.....	59.3	81.8	45.5
Income tax liability.....	79.5	84.9	76.3
Refund.....	70.2	65.5	73.1
Presidential campaign fund contribution.....	34.2	29.5	37.1
Official preaddressed label..	59.3	56.1	61.3
Official IRS envelope.....	79.8	66.3	88.1
IRS mailed package return....	58.1	30.9	74.7

NOTE: Of the 89,154,000 total returns, there were 33,820,000 preparer and 55,333,000 self-preparer returns.

*Dr. Grayson is Chief of the Special Studies Section. He is under the direction of John P. Hiniker, Chief of the Projections and Special Studies Branch.

Table 2.--Percentage Frequency of Selected Forms and Schedules Among Preparer and Self-Preparer Form 1040 Returns, Tax Year 1980

	Form 1040 returns		
	All	Preparer	Self-preparer ¹
	(1)	(2)	(3)
Schedule:			
A.....	51.7	51.9	51.3
B.....	42.5	43.4	41.4
C.....	15.5	19.5	11.2
D.....	12.1	14.6	9.3
E.....	19.3	24.7	13.5
F.....	4.2	5.7	2.5
G.....	9.3	11.7	6.7
Form:			
2106.....	10.1	9.8	10.2
2210.....	5.7	9.4	1.7
2441.....	8.0	7.9	8.1
3468.....	6.5	10.2	2.3
4797.....	2.5	4.4	(¹)
5695.....	8.8	8.7	8.8

¹Less than 0.05 percent.

NOTE: Of the 52,841,000 total Form 1040 returns, there were 27,673,000 preparer and 25,168,000 self-preparer returns. See page 21 for identity of forms and schedules.

When a form or schedule accompanies a self-preparer 1040, it is almost invariably a standard form designed and printed by IRS--99 percent of the time. Not so with preparer returns, where more than one-fourth of the attachments are of the preparer's design, or from some other non-IRS source. The 1980 TPUS, in covering 18 types of attachments, each with at least 400,000 filed, found that preparer returns contributed over 20 million attachments in a non-standard form. Self-preparer returns, in contrast, contributed only about 700,000.

Thus, for several reasons, preparer returns are likely to be more expensive for IRS to process than self-preparer returns: the former are less likely to have the official mailing label, they frequently have more attached forms and schedules, and their attachments are less likely to be in a standard form.

As shown below, preparer returns for 1980, taken as a group, had a median adjusted gross income (AGI) of about \$16,500; the median for self-preparer returns was \$11,600. However, when this comparison is made within Form 1040 and 1040A returns separately, the difference between preparer and self-preparer returns almost vanishes.

Median Adjusted Gross Income, Returns for Tax Year 1980

	Return		
	All	Preparer	Self-preparer
Total.....	\$13,000	\$16,400	\$11,600
Form 1040.....	20,000	19,200	20,800
Form 1040A.....	7,500	7,500	7,500

The explanation of the apparent paradox is both the generally higher AGI of 1040 returns and, as previously noted, the greater frequency of 1040 usage among preparer returns.

The Table 3 data (portrayed in Figure 1) show, for Form 1040A, a mild tendency for the proportion of preparer returns to decline in the higher AGI classes. The line for Form 1040 shows an even stronger tendency to fall as AGI increases--except in the highest class, for \$50,000 and over. For "All Returns," the line moves progressively upward with increasing size of AGI. This is another version of the "paradox" of the medians, and the explanation is essentially the same. Form 1040 filers generally use preparers more heavily than 1040A filers; and, with increasing income, the proportion of returns filed on Form 1040 increases at the expense of the 1040A. Thus in computing the average percentage of all preparer returns, the (high) percent of preparer returns among 1040's gets an ever-increasing weight as AGI increases.

Table 3.--Percent of Returns With Entry for Paid Preparer's Signature by Type of Return and by Size of Adjusted Gross Income, Tax Year 1980

Size of adjusted gross income	Return		
	All	Form 1040	Form 1040A
	(1)	(2)	(3)
Total.....	37.9	52.4	16.9
Under \$5,000.....	27.7	59.7	15.8
\$5,000 under \$10,000.....	33.8	54.0	19.6
\$10,000 under \$15,000.....	34.3	52.8	15.3
\$15,000 under \$20,000.....	41.4	53.9	16.8
\$20,000 under \$30,000.....	45.4	50.0	16.1
\$30,000 under \$50,000.....	47.8	48.8	*14.3
\$50,000 and over.....	55.3	55.3	(¹)

*Estimate should be used with caution because of small number of sample returns on which it is based.

¹No returns in sample.

Table 4 and Figure 2 show that the 1980 "All returns" pattern was basically the same as the average for the entire 1977-1980 period, i.e., higher income taxpayers tended to use preparers more than did lower income taxpayers. The 1975-1976 line suggests that, prior to 1977, taxpayers in the "middle income" levels (\$15,000 to \$30,000) tended to use preparers somewhat less than those with incomes of \$5,000 to \$15,000. The decline in indicated average usage appears concentrated in the income classes below \$15,000.

If most of the indicated decline for 1976-1977 was due to the nonsigning of returns by unpaid preparers after 1976, it may be that unpaid preparers were--and perhaps still are--relatively more important in the lower income classes than in the higher ones.

THE WAGE AND TAX WITHHOLDING STATEMENT, FORM W-2

The Form W-2 is the most important of a small family of forms dealing with tax withheld. It relates to more taxpayers, covers more dollars of income, and there are more of them than any other in the "W" group. (In 1981, about 150 million Forms W-2 were filed with tax returns.) Per square inch, the W-2 probably contains more information than any other tax

Figure 1
Percentage Frequency of Preparer Returns by Adjusted Gross Income Class by Type of Return, Tax Year 1980

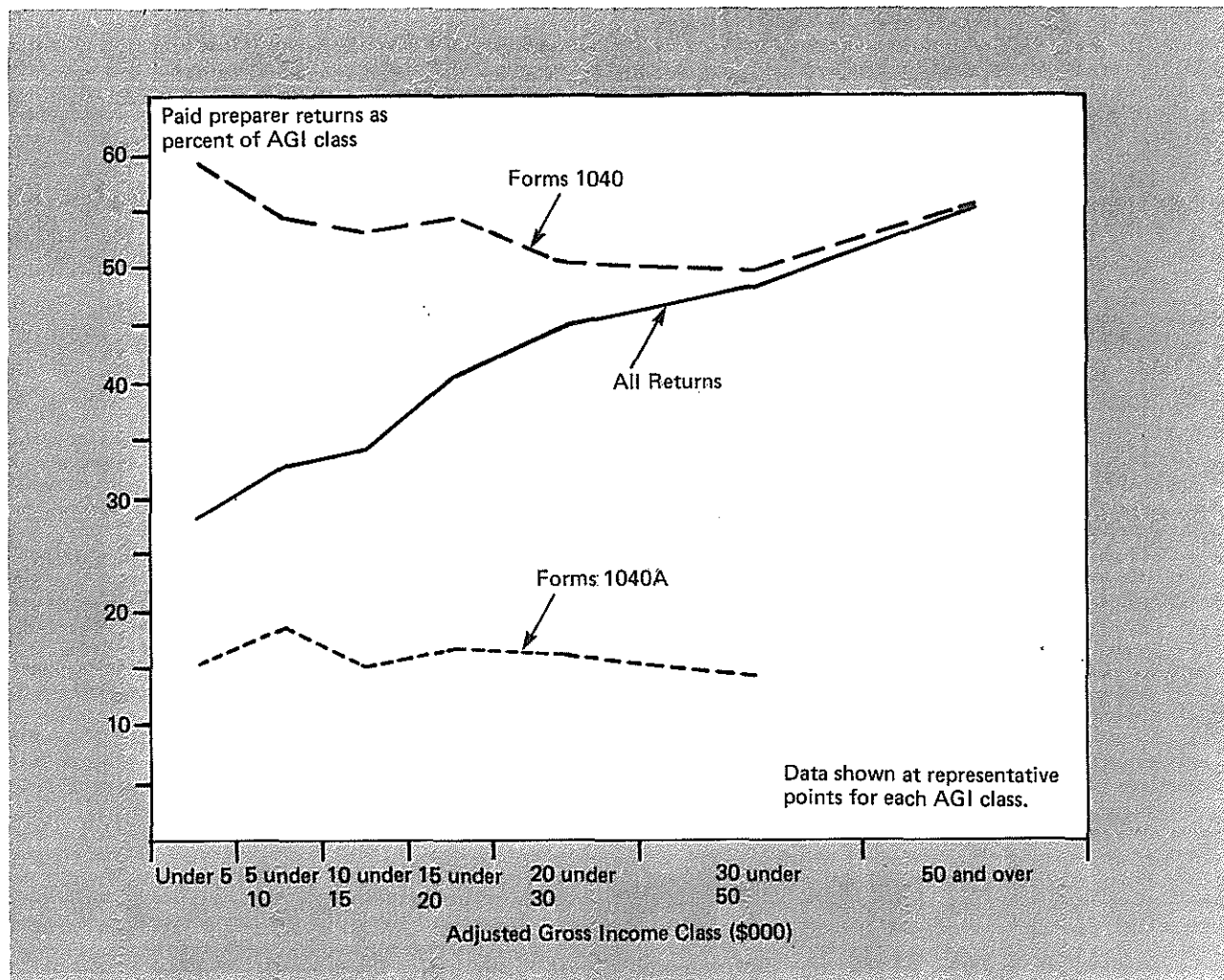


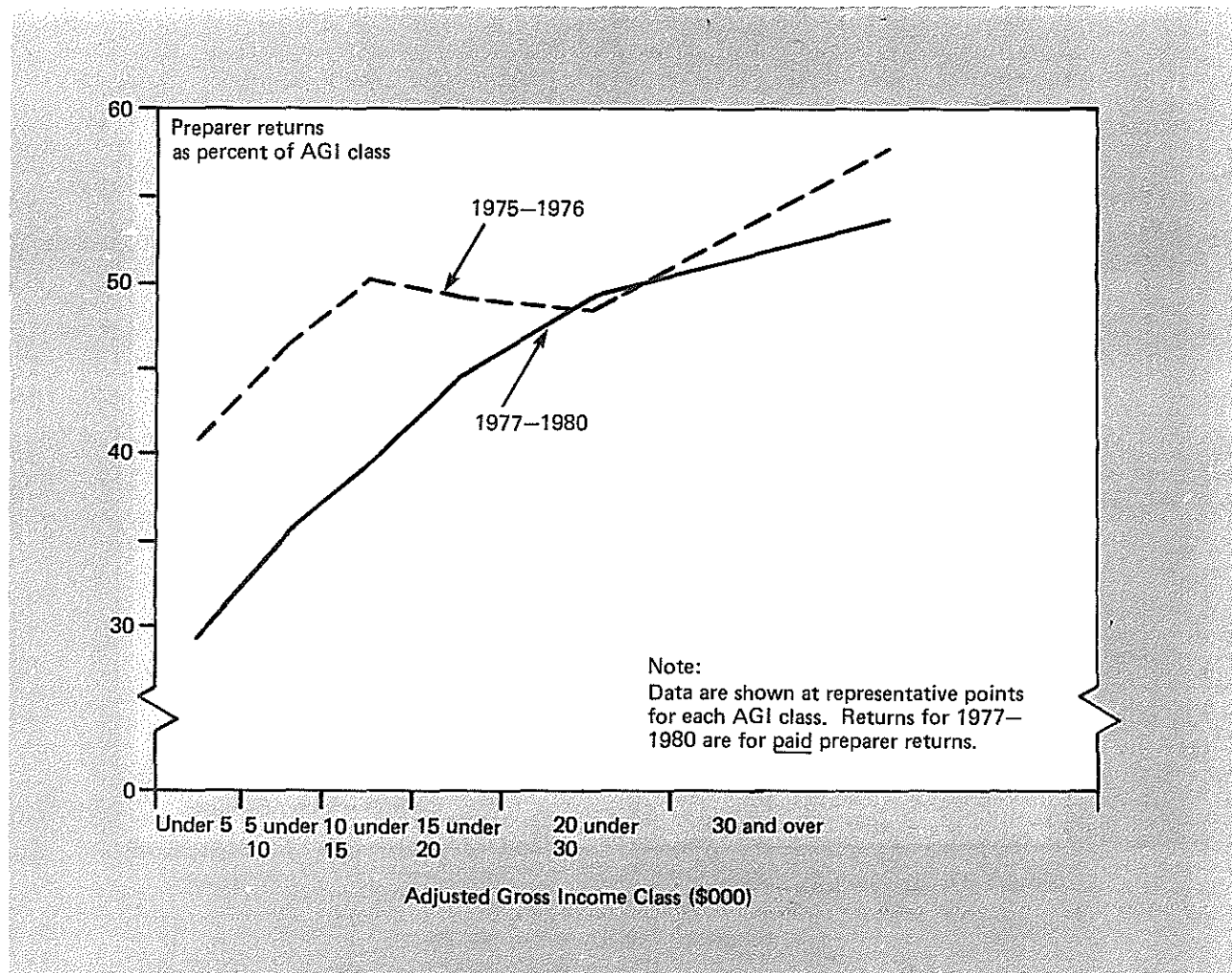
Table 4.--Percent of Returns With Entry for Preparer Signature by Size of Adjusted Gross Income, Tax Years 1969, 1975-1980

Tax year	Average	Size of adjusted gross income					
		Under \$5,000	\$5,000 under \$10,000	\$10,000 under \$15,000	\$15,000 under \$20,000	\$20,000 under \$30,000	\$30,000 and over
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1980.....	37.9	27.7	33.8	34.3	41.4	45.4	49.2
1979.....	39.3	28.8	35.1	39.0	43.3	46.0	52.9
1978.....	39.3	27.7	35.4	39.1	44.6	49.0	54.6
1977.....	42.4	32.7	38.8	45.0	48.2	51.6	55.1
1976.....	47.3	41.0	47.7	52.5	48.8	47.1	56.4
1975.....	45.5	40.1	46.0	47.7	47.5	46.9	58.5
1969.....	52.9	46.0	59.6	55.5		55.5	74.8

NOTE: For 1977-1980: Returns read essentially, "Paid preparer's signature"; for other years, simply "Preparer's signature."

Figure 2

Percentage Frequency of Preparer Returns by Adjusted Gross Income Class Tax Years 1975-76 and 1977-80



form. It tells the taxpayer and IRS the amount of wages paid by an employer subject to Federal income tax and to social security tax (or "FICA"), the amount withheld for each purpose and for State income tax, and provides miscellaneous other information. In addition, it is the only source of information in the files of the tax system on such economic data as single-earner versus two-earner married couples and number of employers of workers in the course of a year. Further, since the W-2 contains the employer identification number (EIN), it provides a link to tax and economic data on the employer in IRS's Business Master File.

For Tax Year 1980, virtually every Form 1040A had at least one W-2 attached, regardless of income class (see Table 5). In contrast, about eight in every ten Forms 1040 had at least one W-2. The proportion of 1040's with W-2's within an income class, however, depends on the level of income, rising from only about one-half in the lowest AGI class to better than 90 percent in the higher classes (but falling to 87 percent in the AGI class \$50,000 and over). Overall, 89 percent of individual income tax returns had one or more attached W-2's.

Over the past 11-year period, this proportion has been measured by TPUS seven times and has never been found to be more than one percentage point above or below 89 percent. At the start of this period, the

average number of W-2's per return filed was just shy of 1.7; in the late 1970's it was close to 1.6 per return.

Tax year	Returns with Form W-2 (Percent)	Average Forms W-2 per return
1970.....	89.5	1.68
1973.....	90.0	1.69
1974.....	88.5	1.65
1975.....	88.4	1.59
1976.....	88.6	1.57
1978.....	88.9	1.62
1980.....	88.9	n.a.

As might be expected, there are substantial differences between joint and nonjoint returns in W-2 reporting. Based on 1976 TPUS information, almost three-fifths of nonjoint returns had only one attached W-2, compared with only about one-third of the joint returns (see Table 6). Conversely, almost another third of the joint returns, but less than 20 percent of nonjoint returns, had two W-2's. Table 6 also shows that Forms 1040 and 1040A differ mainly with respect to the zero and one W-2 categories.

Table 5.--Percent of Returns With Form W-2 by Type of Return and by Size of Adjusted Gross Income, Tax Year 1980

Size of adjusted gross income	All returns	Form 1040 returns	Form 1040A returns
	(1)	(2)	(3)
Total.....	88.9	82.4	98.5
Under \$5,000.....	85.0	51.0	97.6
\$5,000 under \$10,000.....	84.2	63.4	98.8
\$10,000 under \$15,000.....	88.8	79.4	98.4
\$15,000 under \$20,000.....	91.3	87.1	99.7
\$20,000 under \$30,000.....	94.1	93.2	100.0
\$30,000 under \$50,000.....	94.1	93.9	100.0
\$50,000 and over.....	87.0	87.0	(1)

¹No returns in sample.

Table 6.--Percentage Distribution of Returns by Number of Attached Forms W-2 by Filing Status; by Return Type, Tax Year 1976

Number of Forms W-2	Total	Filing status		Type of return	
		Joint	Non-joint	Form 1040	Form 1040A
	(1)	(2)	(3)	(4)	(5)
Total.....	100.0	100.0	100.0	100.0	100.0
0.....	11.8	11.8	11.8	18.2	1.8
1.....	46.0	34.6	58.7	39.6	58.7
2.....	26.0	32.6	18.7	26.7	24.2
3.....	10.4	12.9	7.7	10.2	10.4
4.....	3.5	4.8	2.1	3.4	3.0
5 or more.....	2.2	3.4	0.9	2.0	2.0
Average per return..	1.57	1.77	1.33	1.50	1.63

Tables 7 and 8 focus on joint returns. In Table 7 we note, for example, that about 17 percent of the joint tax returns had no W-2 for the husband; about one-half had no W-2 for the wife. We also note that about 43 percent of the joint returns had two earners reported (for 1970, the comparable figure was 41 percent). Table 8, for jointly filed Forms 1040A, shows the steady increase in the proportion of dual-earner returns as income increases. In this case, dual-earner status is presumably a prime determinant of higher income.

CARRYING OUT THE STUDY

In December of each year, instructions are sent to all 10 Internal Revenue service centers (where all income tax returns are filed) to select a systematic sample of 1/13,000th of all mail items received between the following January 1 and June 30. (The bulk of the selection is made through the use of a counting device on the automatic envelope opening machines.) If the selected item contains a Form 1040 or 1040A, the entire contents are reproduced, including the envelope. If it does not include a 1040 or 1040A, only page 1 of the contents is reproduced. On Friday, the accumulated weekly sample is express-mailed to the IRS Statistics Division for Monday or Tuesday delivery. Weekly reports on key items on Forms 1040 and 1040A are issued by Thursday.

Table 7.--Percentage Distribution of Joint Returns by Number of Forms W-2 for Husband by Number for Wife, Tax Year 1976

Number of Forms W-2 for wife	Total	Number of Forms W-2 for husband		
		0	1	2 or more
	(1)	(2)	(3)	(4)
Total.....	100.0	17.4	60.3	22.3
0.....	51.4	11.8	29.9	9.7
1.....	39.5	4.7	25.5	9.3
2 or more.....	9.1	0.9	4.9	3.3

Table 8.--Percent of Jointly Filed Form 1040A Returns by Number of Spouses With Form W-2 and by Size of Adjusted Gross Income, Tax Year 1980

Item	Total	Size of adjusted gross income				
		Under \$5,000	\$5,000 under \$10,000	\$10,000 under \$15,000	\$15,000 under \$20,000	\$20,000 and over
	(1)	(2)	(3)	(4)	(5)	(6)
All joint Forms 1040A filed.....	100.0	100.0	100.0	100.0	100.0	100.0
At least one spouse indicated on Form(s) W-2.....	99.6	98.9	99.4	99.4	100.0	100.0
One spouse only.....	52.3	73.9	72.1	53.9	43.0	34.9
Two spouses.....	47.3	25.0	27.3	45.5	57.0	65.1
No Form W-2 attached.....	*0.4	*1.1	*0.6	*0.6	-	-

*Estimate should be used with caution because of the small number of sample returns on which it is based.

Uncorrected data--as entered on the tax return--are manually abstracted to checksheets and later transcribed into magnetic form; both operations are 100 percent verified. The magnetic data file is tested and corrected to eliminate processing error. The major report is based on sampling through about May 1, so that the large influx of tax returns filed around April 15 can be covered. By the May 1 cutoff, the TPUS sample consists of a total of about 7,500 current-year returns filed on Forms 1040 and 1040A. (Prior-year returns are considered "out of scope.") Sample returns received between May 1 and June 30--about 300 in number--are separately tabulated; these will be later referred to as the "Late Sample." Returns in the "regular" and late samples are permanently retained.

The designed sampling rate has been set at 1/13,000 to yield a sample of about 7,000. However, in most recent years the average effective sampling rate has exceeded the designated rate; in addition, the effective rate for 1040A's has exceeded that for the 1040's. Research to date on the responsible factors has been inconclusive.

The estimation procedure is based on using the inverse of the effective sampling rates for Forms 1040 and 1040A weighted separately. For example, in 1981, the official IRS count of Forms 1040 received by May 1 was 52,841,000, while the number of sample 1040's was 4,367. Thus the effective sampling rate for 1040's was 1/12,100 and the weight assigned was 12,100. For Forms 1040A, the weight was 11,448.

Because the TPUS samples are unstratified, data from them are fundamentally of one kind--frequency of occurrence, generally presented both in absolute numbers at the population level, and as percentage frequencies. Items selected for tabulation predominantly deal with the presence of line item entries and types of attached forms and schedules.

LIMITATIONS OF TPUS

As previously indicated, the TPUS sample--the whole TPUS program, in fact--was designed with a clearly defined objective in mind: to provide "characteristic-type" data and to provide them as early in the filing year as possible. The sample is therefore small, has a uniform designated sampling rate, and is not stratified in the customary sense. (From one point of view, the TPUS sample might be termed stratified by service center and week of receipt of the tax return. While more research might be done in this area, the relative efficiency of these strata with respect to reducing the sampling variability of most items is likely to be quite insignificant.)

The characteristics of the TPUS sample have been important in a variety of ways. Because of the sample's relatively small size and lack of stratification, calculated sampling errors are larger than those experienced in the much larger Statistics of Income sample of individual income tax returns--which has ranged to 150,000 returns or more in the recent past--and is highly stratified to boot. IRS administrators and others have, nevertheless, made extensive use of the TPUS reports for characteristic-type data.

Information on sampling error assists in the interpretation of TPUS results by enabling the user to set confidence limits with known probability on the estimates. From the illustration given below, if a characteristic occurred on 50 percent of the approximately 50 million Forms 1040 in the Tax Year 1980 TPUS, it had a relative sampling error, or coefficient of variation (CV), of 1.6 percent. Now 1.6 percent of 50 is 0.8 percent; and 50 ± 0.8 is 49.2 and 50.8. One can, therefore, have 68 percent confidence that the percentage in the population is between 49.2 and 50.8 percent. For 95 percent confidence, one doubles the CV to obtain the range, 48.4 to 51.6 percent.

Estimated percentage with characteristic	Size of population (with and without characteristic)-- Coefficient of variation (Percent)	
	20,000,000	50,000,000
1.5	19.9	12.5
10.0	7.4	4.7
50.0	2.5	1.6

The sampling error of dollar estimates derived from an unstratified sample is a matter of much greater concern, when high tax or income returns as well as low are being sampled at an across-the-board rate of 1/13,000. The original practice has therefore been retained of not using TPUS to estimate dollar amounts, except for very special and limited purposes.

The TPUS sample, by its very nature, is an incomplete sample. Past experience indicates that the regular TPUS report has covered about 95 percent of the total individual returns filed each year. TPUS counts of returns with a given characteristic are bound to be smaller than the count for the entire year. TPUS percentages may, or may not, differ from the population for the year, depending on whether the variable being measured is present in the same proportion among returns received before and after May 1. Items vary in this respect, our analyses show (see the section on the Late Sample, below). Further experimentation with ratio estimating techniques may lead to new approaches to presenting the sample data, both with respect to frequencies and dollar amounts. (See, for example, Peter Sailer and Noreen Hoffmeier: "Early Highlights from 1980 Individual Income Tax Returns," SOI Bulletin 1(2), Fall 1981, pp. 1-5.)

Rare items--or rare combinations of items--can be measured by the TPUS sample, but only with much imprecision, as already suggested. For example, the coefficient of variation (CV) for an estimate of 200,000 returns approaches 25 percent. (At the current sampling rate, an estimate of 200,000 is based on only 17 sample returns.) The CV's for smaller estimates are even larger. A characteristic that occurred on only 10,000 tax returns in the population might well not appear in the sample at all. TPUS reports, where necessary, therefore may combine classes or may indicate that the datum is either missing in the sample or possesses a sampling error too high for it to be considered reliable.

THE REPRESENTATIVENESS OF TPUS--A LOOK AT THE RECORD

Viewed against the background thus far indicated, it is reasonable to ask, "How good are the TPUS statistics? How does the total TPUS system, with its limitations, stack up against an accepted benchmark?" The most useful benchmark, because of its wealth of detail, is the annual series, Statistics of Income--Individual Income Tax Returns (SOI).

The upper half of Table 9 shows how the TPUS distributions for AGI in four recent years closely tracked those of SOI. The lower half shows the value of "d," the relative difference between the TPUS and SOI estimates in each AGI class/year cell, as a percentage of the SOI estimate. The overall average for 24 cells (irrespective of sign) is 3.8 percent. (An average coefficient of variation for the TPUS estimated percentages is about 2.5 percent.)

In all four years, TPUS underestimated the percentage importance of the under \$5,000 class and the class of \$30,000 and over. No valid reason for the former phenomenon is apparent. The situation in the high income class, however, is consistent with the greater frequency of high income returns among returns filed after May 1 than before that date (see section,

Table 9.--Percentage Distribution of Tax Returns by Size of Adjusted Gross Income, Tax Years 1976-1979--Taxpayer Usage Study and Statistics of Income Compared

Tax year, source	Total or average ¹	Size of adjusted gross income					
		Under \$5,000	\$5,000 under \$10,000	\$10,000 under \$15,000	\$15,000 under \$20,000	\$20,000 under \$30,000	\$30,000 and over
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1976:							
TPUS.....	100.0	27.0	23.9	17.9	13.7	12.4	5.1
SOI.....	100.0	28.3	23.5	17.2	13.2	12.2	5.6
1977:							
TPUS.....	100.0	26.3	22.9	16.5	13.4	14.0	7.0
SOI.....	100.0	26.9	22.3	16.5	13.2	14.0	7.2
1978:							
TPUS.....	100.0	23.1	22.4	² 17.0	12.8	16.0	8.7
SOI.....	100.0	24.2	21.9	15.8	12.7	15.5	9.9
1979:							
TPUS.....	100.0	21.2	21.9	16.3	² 13.1	16.4	11.1
SOI.....	100.0	22.5	21.2	15.6	12.3	16.6	11.9
Percent difference ³ (d) in distribution percentage = 100 (TPUS-SOI) / SOI							
1976.....	4.0	-4.6	1.7	4.1	3.4	1.6	-8.9
1977.....	1.5	-2.2	2.7	0.0	1.5	0.0	-2.8
1978.....	5.1	-4.5	2.3	7.6	0.8	3.2	-12.1
1979.....	4.7	-5.8	3.3	4.5	6.5	-1.2	-6.7
Average ¹	3.8	4.3	2.5	4.0	3.0	1.5	7.6

¹Average = $\sum |d| \div n$ where d is irrespective of sign.

²Direction of change from previous year in TPUS not the same as in SOI.

³Example: For 1976, Under \$5,000: $100 (27.0 - 28.3)/28.3 = -4.6$ percent.

below). While these high income late returns are sampled for SOI, they are not sampled for the regular TPUS report.

Whether an AGI class is gaining or losing in relative importance from year to year is indicated almost as well by TPUS as by SOI. Thus TPUS showed the class below \$5,000 with a drop of 0.7 percentage point from 27.0 percent in 1976 to 26.3 in 1977. SOI also showed a 1976/1977 decline. Out of 18 such comparisons of the two sources for year-to-year change, all but two (noted in the top half of the table) showed TPUS moving in the same direction as

SOI. (Comparison of the SOI preliminary data for 1980 --published in the preceding article--with the 1980 TPUS provides similar results in terms of levels and direction of change.)

Amounts of change in the percentage importance of the AGI classes can also be derived from the top half of the table. In the previous example, a 0.7 percentage point drop was noted from 1976 to 1977 in TPUS for the lowest income class. The total of such changes over the four-year period for the two sources, by income class, is as follows:

Source	Sum of percentage changes					
	Under \$5,000	\$5,000 under \$10,000	\$10,000 under \$15,000	\$15,000 under \$20,000	\$20,000 under \$30,000	\$30,000 and over
	(1)	(2)	(3)	(4)	(5)	(6)
TPUS.....	-5.8	-2.0	-1.6	-0.6	4.0	6.0
SOI.....	-5.8	-2.3	-1.6	-0.9	4.4	6.3

These are rather remarkable results, comparable to an engineering structure that resists much higher stresses than it was designed to withstand. There are limits, however, to the "strength" of the TPUS system, and the primary ones must be borne in mind: the small sample size and the concentration on the first four months of receipts.

Both the strength and weakness of the total TPUS program are further revealed by a consideration of the counts of certain characteristics published in the TPUS report. One of the major areas covered in each report is the counts of Form 1040 returns with specified forms and schedules attached. Historically, TPUS counts have fallen short of SOI indications. (Since SOI does not count, for example, "returns with Schedule A," the SOI count of "returns with itemized deductions" can be taken as the equivalent.)

The accompanying table of Tax Year 1978 data (Table 10) supports the basic soundness of TPUS sampling and processing, but also indicates the limitations placed on the program by the need for a report based on an early May cut-off. About 4 million returns, represented by about 300 late TPUS sample returns, are filed in May and June. About one or two million more returns are received in the rest of the year. Taking Schedule A as an example, one notes that the regular TPUS reported 23.8 million returns with this schedule, or 27.5 percent of all returns covered by the regular sample. With the addition of the late sample, the number of returns with Schedule A becomes 25.7 million (compared with SOI's 25.8 million) and comprises 28.4 percent of the augmented population (compared with SOI's 28.8 percent).

The data from the regular TPUS on the presence of the indicated schedules, in other words, is entirely reasonable for the period covered by the report. When the regular TPUS counts and percentages are supplemented by late sample data, the combined results are virtually on a par with SOI.

Comparing the Late Sample to the Regular Sample

The late sample of returns filed during May and June is characterized by somewhat higher adjusted gross income compared with the earlier returns filed. Table 11 indicates that incomes below \$10,000 are less common in the late sample, while incomes over \$50,000 are more frequent.

Table 10 has already shown (see the second and fourth columns) that the late sample, compared with the regular sample, is much more prone to have attached schedules such as Schedules A, C, D, and G. Close to one-half of the late sample, for example, had an attached Schedule A compared with only a bit more than one-fourth of the regular sample.

Two features that further contrast late and regular TPUS samples (see Table 12) are the preponderance of the Form 1040 and the greater frequency of the signature of paid preparers among the late sample returns. Undoubtedly associated with these characteristics are the significantly lower levels of usage in the late sample of official IRS label, official envelope, and mailed package return. Requests for refund, as might be expected, were only 76 percent as frequent in the 1979 late sample as in the regular sample, but entries for income tax occurred about equally often in the two groups.

Returns received after April, then, are likely to be more expensive per unit to process; they are more likely to have attached schedules and thus be more complex, to have higher income, and to be less likely to be associated with the standardized forms, labels and envelopes that would simplify IRS processing.

Table 10.--Returns With Five Selected Schedules, Tax Year 1978--Taxpayer Usage Study and Statistics of Income Compared

[Numbers of returns in millions]

Schedule	TPUS						SOI	
	Regular		Late sample		Combined		Number	Percent with schedules
	Number	Percent with schedules	Number	Percent with schedules	Number	Percent with schedules		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total returns..	86.4	-	4.2	-	90.6	-	89.7	-
Schedule:								
A.....	23.8	27.6	1.9	45.8	25.7	28.4	25.8	28.8
C.....	7.5	8.7	1.3	30.9	8.8	9.7	8.2	9.1
D.....	7.4	8.6	0.9	21.0	8.3	9.2	8.6	9.6
F.....	2.6	3.0	0.2	4.8	2.8	3.1	2.9	3.0
G.....	3.5	4.0	0.6	15.3	4.1	4.6	4.5	5.0

Table 11.--Size of Adjusted Gross Income Percentage Distribution, Regular TPUS versus Late Sample, Tax Year 1979

Source	Total	Size of adjusted gross income						
		Under \$5,000	\$5,000 under \$10,000	\$10,000 under \$15,000	\$15,000 under \$20,000	\$20,000 under \$30,000	\$30,000 under \$50,000	\$50,000 and over
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Late.....	100.0	17.2	15.7	17.9	13.8	17.4	10.4	7.5
Regular.....	100.0	21.2	21.9	16.3	13.1	16.4	9.2	2.0

Table 12.--Presence of Selected Characteristics, Regular TPUS versus Late Sample, Tax Year 1979

Source	Percent of source with characteristic								
	Form 1040	Preparer signature	Official IRS label	Official envelope	Mailed package return	Computer printed	Entry for--		
							Income tax	Refund	Earned income credit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Late.....	81.8	64.5	46.1	51.8	41.3	3.1	81.1	55.0	5.9
Regular.....	58.2	39.3	59.6	76.8	59.2	3.3	78.5	72.5	6.9

NOTE TO TABLE 2

IDENTITY OF LISTED SCHEDULES AND FORMS

<u>Schedule</u>	<u>Title</u>	<u>Form</u>	<u>Title</u>
A	Itemized Deductions	2106	Employee Business Expenses
B	Interest and Dividend Income	2210	Underpayment of Estimated Tax by Individuals
C	Profit or (Loss) from Business or Profession	2441	Credit for Child and Dependent Care Expenses
D	Capital Gains and Losses		
E	Supplemental Income Schedule	3468	Computation of Investment Credit
F	Farm Income and Expenses	4797	Supplemental Schedule of Gains and Losses
G	Income Averaging	5695	Residential Energy Credit