ALTERNATIVE DESIGNS FOR A CROSS-SECTIONAL SAMPLE OF INDIVIDUAL TAX RETURNS: THE OLD AND THE NEW

Allen L. Schirm and John L. Czajka, Mathematica Policy Research, Inc.
Allen L. Schirm, 600 Maryland Avenue, S.W., Suite 550, Washington, D.C. 20024-2512

1. INTRODUCTION

The Statistics of Income (SOI) Division of the Internal Revenue Service (IRS) has undertaken a major redesign of its annual sample of individual tax returns. The SOI cross-sectional sample supports two main statistical activities: (1) development of aggregate estimates of income and tax components by SOI and the Bureau of Economic Analysis (BEA) and (2) tax policy modeling by the Office of Tax Analysis (OTA) in the Department of Treasury and the Joint Committee on Taxation of the United States Congress. The principal objective of the redesign of the cross-sectional sample is to enhance the sample's usefulness for tax policy modeling without unduly compromising the sample's ability to support precise aggregate estimates.[1]

This paper has four principal objectives:

1. to describe and compare the old and new sample designs

2. to examine how returns are treated differently under the old and new designs

3. to compare the abilities of the old and new designs to support precise aggregate estimates of income and tax components

4. to compare the abilities of the old and new designs to support tax policy modeling

The paper uses data from the 1987 and 1988 SOI samples to pursue the latter three objectives.

The next section of this paper describes and compares the income definitions and strata definitions of the old and new sample designs. Section 3 examines how returns sampled under the old design would be stratified under the new design. Section 4 compares the precision of aggregate estimates under the old and new designs, and Section 5 compares sample counts under the two designs. Section 6 summarizes the results obtained.

2. THE OLD AND NEW SAMPLE DESIGNS

Samples drawn under the new design, like those drawn under the old design, are highly stratified by income. Income, however, is defined differently under the two designs. Moreover, the two designs treat returns differently depending on, for example, the composition of income and the presence of certain forms or schedules attached to the tax return. The new design seeks to identify and sample at higher rates returns that are most useful for policy modeling.[2] This section describes differences between the two designs in defining income and strata.

2.1 Income Definitions

The design dubbed "old" in this paper was first used to draw a 1982 SOI sample from the population of returns filed in 1983 (for, primarily, tax year 1982). For stratifying returns, the principal definition of income under the old design is the larger absolute value of a positive amounts total (PAT) and a negative amounts total (NAT), calculated from the income components of a taxpayer's adjusted gross income (AGI). The old design selects a relatively small number of returns on the basis of combined business and farm total receipts rather than PAT/NAT. The components of PAT/NAT have changed slightly over time with changes in the tax form.

The PAT components used to select the 1987 SOI sample were: salaries and wages, taxable interest income, dividends, taxable refunds of state and local income taxes, alimony received, net business profits (if positive), net capital gains, capital gains distributions reported on Form 1040, net Form 4797 income (if positive), taxable pensions/annuities, gross rent/royalty income, net farm rental income (if positive), gross partnership/S-corporation income, gross estate/trust income, Schedule E windfall profits tax refund, net farm profits (if positive), unemployment compensation, taxable social security income, and net other income (if positive). Beginning in 1988, taxable IRA distributions were no longer included in pension and annuity income and were, instead, a separate component of AGI and PAT.

The NAT components used to select the 1987 SOI sample were: net business profits (if negative), net capital losses, net Form 4797 income (if negative), gross allowable rent/royalty losses, net farm rental income (if negative), gross allowable partnership/S-corporation losses, gross allowable estate/trust income, Schedule E windfall profits tax deduction, net farm profits (if negative), and net other income (if negative).

The new sample design, which is being used for the first time to draw a 1990 SOI sample, assigns returns to strata primarily according to either total gross positive income or total gross negative income depending on whether total net income is nonnegative or negative, respectively. Total net income is the sum of total gross positive income and total gross negative income, both of which are obtained by summing various individual income components.

The components of gross positive income are: salaries and wages, total interest income, dividends, alimony received, gross business profits (if positive), short-term capital gains, long-term capital gains less gain from sale of home, capital gains distributions reported on Form 1040, net Form 4797 income (if positive), tested pension and annuity income, taxable IRA distributions, gross rent/royalty income, net farm rental income (if positive), gross partnership/S-corporation income, gross trust/estate income, gross farm income (if positive), unemployment compensation, tested social security income, and net other income (if positive).[3,4]

The components of gross negative income are: gross business profits (if negative), total business deductions, short-term capital losses, long-term capital losses, net Form 4797 income (if negative), gross allowable rent/royalty losses, net farm rental income (if negative), gross allowable partnership/S-corporation losses, gross allowable trust/estate losses, gross farm income (if negative), total farm expenses, net other income (if negative), alimony paid, and moving expenses.

The key differences between the two designs' income definitions are:

*   The new design includes in total income nontaxable amounts for some items (interest income, pension and annuity income, and social security income).[5]

*   The new design uses net income amounts only when gross amounts are unavailable at the time of sample selection.[6]

The latter is especially important. Under the old design, large gross business profits, gross farm profits, or gross capital gains can be offset entirely or at least partly by, respectively, business deductions, farm expenses, or gross capital losses. Under the new design, positive and negative amounts are separated to the greatest extent possible. A return with $750,000 in business gross profits, $800,000 in business expenses, and no other income/gain or expense/loss amounts would be stratified on the basis of a gross negative income of $800,000 rather than a NAT of $50,000.

2.2 Strata Definitions

With respect to strata definitions, the key changes that will occur when the new design replaces the old design are:

- the near elimination of hierarchical stratum assignments

- the addition of negative income strata, allowing the separation of returns stratified on the basis of negative amounts from returns stratified on the basis of positive amounts

- the addition of strata for returns that are more useful for policy modeling, allowing higher sampling rates for those returns and lower sampling rates for returns of equal total income that are less useful for policy modeling

Table 1 gives 1987 sample counts and sampling rates for the new design strata, the former obtained by applying the latter to the filing population from which the 1987 SOI sample was drawn.[7] A comparable table pertaining to the old design is available upon request. Hostetter et al. (1990) describe the new sample design in greater detail.

Hierarchical Stratum Assignments. Under the old design, sample selection follows a strict hierarchy. A return is first subject to selection as a high income, nontaxable (HINT) return. A return that is not a HINT is then subject to selection as a return with high combined business (Schedule C) net profit/loss. These first two strata are always certainty strata.[8] If a return is not selected in one of these first two certainty strata, it is assigned to a stratum based on the forms or schedules attached to the return (and on income) according to the following hierarchy: (1) Form 2555 (Foreign Earned Income) attached (five income strata); (2) Form 1116 (Computation of Foreign Tax Credit) but no Form 2555 attached (five income strata); (3) Schedule C (Profit or (Loss) From Business or Profession) but no Form 2555 or Form 1116 attached (nine income strata); (4) Schedule F (Farm Income and Expenses) but no Form 2555, Form 1116, or Schedule C attached (nine income strata); and (5) no Form 2555, Form 1116, Schedule C, or Schedule F attached (nine income strata).

The principal implication of this hierarchical sample design is that, with little regard for the returns' usefulness for policy modeling, two returns with equal incomes could be subject to very different selection probabilities simply depending on what forms or schedules are attached to the returns. In 1987, a return with total income (PAT/NAT) of $3,000,000 would have been selected with certainty if it had a Form 2555, with probability 25% if it had a Form 1116 but no Form 2555, and with probability 50% if it had a Schedule C but no Form 2555 or Form 1116, for example. Moreover, even if returns with certain forms or schedules attached are regarded as more useful for policy modeling, there is no assurance that a return with a Form 1116 or a Schedule C attached, for instance, will be sampled at a higher rate than a return with equal PAT/NAT but no Form 1116 or Schedule C attached. Sampling rates for strata in different levels of the hierarchy are not constrained to satisfy any particular relationships.

Under the new sample design, hierarchical stratum assignment is nearly eliminated. Before a return is assigned to 1 of 24 strata based on total gross positive or negative income, it may be assigned to 1 of 2 special strata. The first consists of HINTs. The second consists of returns with high combined business and farm total receipts. A threshold of $50,000,000 will be used for at least the first year of sample selection under the new design. It is expected that about 50 returns will be selected in this second special stratum. Both strata are certainty strata.

After HINT and high total receipts returns are selected, the new design stratifies the remaining returns entirely on the basis of total gross positive income or total gross negative income. The only exception is for low gross positive income returns, which are distinguished--within income ranges--according to their usefulness for policy modeling. The new design does allow for selecting additional returns of certain types, should that ever become desirable.[9]

Negative Income Strata. Under the old sample design, there are no separate strata for returns with NAT exceeding PAT. Returns with NAT between $50,000 and $100,000 (in absolute value), for example, fall in the same stratum as returns with PAT between $50,000 and $100,000, if the same forms and schedules are attached.

Under the new sample design, returns with negative total net income are assigned to 1 of 9 strata on the basis of total gross negative income. Returns with positive or zero total net income are assigned to 1 of 15 strata on the basis of total gross positive income.

Separating positive and negative income strata reduces within-stratum heterogeneity and increases the precision of aggregate estimates. It also allows higher sampling rates for negative income returns, all of which are regarded as useful for policy modeling (although not equally useful). The new design constrains the sampling rate for a negative income return to be at least as high as any positive income return with equal income in absolute value.

Stratification and Policy Modeling. Under the old sample design, returns with equal total income are treated differently only if they have different forms (2555 or 1116) or schedules (C or F) attached (unless one return is a HINT or high business net profit/loss return or is stratified on the basis of total receipts). Stratum assignments do not depend on the income amounts appearing on the forms or schedules, except insofar as the amounts contribute to PAT/NAT. Moreover, stratum assignments do not depend on the presence or amount of income from capital gains, partnerships, retirement income, and other such sources, except, again, for the contributions to PAT/NAT.

The new sample design seeks to identify systematically returns that will be more useful for policy modeling. At the lower levels of total gross positive income, returns are distinguished as more or less useful for policy modeling and placed in different strata depending on the presence of certain filing characteristics, the composition of total gross positive income, and the level of total gross negative income relative to total gross positive income. This is shown in Table 1. Returns with substantial investment income relative to total income and returns with substantial shares of both labor income and retirement income, for example, are judged more useful for policy modeling.[10] Separately stratifying returns with equal total income but varying usefulness for policy modeling allows more useful returns to be sampled at higher rates and less useful returns to be sampled at lower rates. Members of the redesign team believed that no useful distinctions can be made among returns with total gross income above a certain level ($250,000).

3. A COMPARISON OF STRATUM ASSIGNMENTS UNDER THE OLD AND NEW DESIGNS

For returns in the 1987 and 1988 SOI samples, it is easily shown that, as expected, returns in any given old design stratum typically fall in several new design strata. The returns selected on the basis of high business net profits/losses under the old design are particularly heterogeneous, falling in all new design strata except the two (Strata 10 and 25) from which such returns are excluded by definition. Tables cross-tabulating the old and new design stratum assignments of returns in the 1987 and 1988 SOI samples are available upon request.

Had a 1987 sample been drawn under the new design, the sampling rates in Table 1 would have been used. By determining the new design stratum assignment of each return drawn under the old design, it is possible to calculate, for each new design stratum, the sampling rate that is implied by the old design. The actual and implicit sampling rates can be compared to see what returns are sampled under one design but not the other.

Relative to the new sample design, the old design undersamples returns in all of the new design's negative income strata (Strata 1-9) and high positive income ($1,000,000+) strata (Strata 21-24). The old design also undersamples returns in the low positive income strata containing returns judged more useful

for policy modeling (Strata 12, 14, 16, and 18). Across these latter strata, the old design captures only about 63% as many returns as the new design. Across the negative income strata, the old design captures about 71% as many returns as the new design. Across the high positive income strata, the old design captures about 77% as many returns as the new design. To obtain nearly the same total sample size, the old design captures over half again as many returns in the low positive income strata containing returns judged less useful for policy modeling (Strata 10, 11, 13, 15, and 17). The old design also captures about 7% more returns in the $500,000 to $1,000,000 gross positive income stratum (Stratum 20).

The figures just reported pertain to 1987. However, even when the old design samples over 15,000 more returns than the new design target of 95,000 returns, as in 1988, the old design undersamples negative income returns, high positive income returns, and low positive income returns judged more useful for policy modeling. In these categories, the old design captures only 77%, 87%, and 73% as many returns, respectively, as the new design.[11]

4. PRECISION OF AGGREGATE ESTIMATES UNDER THE OLD AND NEW DESIGNS

Table 2 gives coefficients of variation of sample mean estimates for selected income and tax components. As noted earlier, the principal objective of the cross-sectional sample redesign is to enhance the usefulness of the sample for policy modeling without sacrificing too much precision in aggregate estimates derived from the sample. For most income and tax components, the optimal (Neyman) allocation for estimating a sample mean as precisely as possible would select heavily from low income strata. The returns obtained, however, would not support analyses of most proposed policy initiatives. Relatively few returns with capital gains or losses, for example, would be available.

According to Table 2, implementation of the new design leads to a loss of precision in 1987 for only a few income and tax items, most of which are estimated precisely under both designs. For most income and tax items, the new design supports more precise estimates, and the gains in precision are often substantial. The coefficients of variation for tax-exempt interest, gross short-term capital losses, gross long-term capital losses, gross rent/royalty income, and gross rent/royalty losses are reduced by over 50%. Generally, the old design tends to support slightly more precise estimates for net income items and a few items closely associated with AGI, such as total tax liability. The new design tends to support more precise estimates for most income and tax items, and the gains in precision are especially large for gross income items. For 1987, some of these gains in precision from the new design are attributable to its larger sample size (by nearly 1,700 returns).[12]

5. SAMPLE COUNTS UNDER THE OLD AND NEW DESIGNS

The principal objective of the SOI sample redesign is to enhance the sample's ability to support policy modeling. As discussed earlier, this enhancement is accomplished by improving the mix of sampled returns so that the sample better provides items and dollar amounts that might be relevant to assessing possible policy initiatives. Even with a highly stratified design, many simple low income returns are selected.[13] Thus, the mix of sampled returns is improved by defining income and strata and by specifying sampling rates so as to select more returns with such unusual characteristics as long-term capital gains, rent/royalty income, and tax credits.

Table 3 displays sample counts, by AGI class, obtained under the old and new designs. All sample counts reported pertain to 1987. The substantial difference in new design and old design sample sizes for 1988 makes comparisons of sample counts difficult. When we compare the rates at which certain types of returns--returns with tax-exempt interest, for example--are obtained, we observe similar patterns for 1987 and 1988.

As expected, the new SOI sample design selects more returns from both tails of the AGI distribution, although the difference in the right tail is small. According to Table 3, the new design yields 27% more returns with deficits and 3% more returns with AGI exceeding $250,000. Also as expected, the new sample design selects more returns with unusual items, especially at lower total income (AGI) levels. The new design has 23% more returns with tax-exempt interest income than the old design, 47% more (629 returns vs. 429 returns) in the $0 to $30,000 AGI range. The new design has 21% more returns with gross long-term capital gains and 26% more returns with gross long-term capital losses. The sample count increases in the $0 to $30,000 AGI range are 71% and 100% for gross long-term capital gains and gross long-term capital losses. The new design offers 25% more returns with gross rent/royalty losses and 13% more returns with gross partnership/S-corporation losses. In the $0 to $30,000 AGI range, the new design has 85% more returns with gross rent/royalty losses and 73% more returns (1,166 returns vs. 675 returns) with gross partnership/S-corporation losses. Compared to a sample drawn under the old design, a sample drawn under the new design has more returns with statutory adjustments (5% more), itemized deductions (6% more), and tax credits (21% more) and more returns owing the alternative minimum tax (46% more).

Although the new design selects 21% more returns claiming at least one tax credit, it selects 17% fewer returns claiming a child care credit, 12% fewer in the under $30,000 AGI range. The new design also obtains 8% fewer head of household returns. Thus, the new design may provide a somewhat weaker basis for analyzing tax policies concerning child care, although it does select 11% more returns with dependent children at home and low income (AGI under $30,000). The new design selects 8% fewer returns with income from unemployment compensation, 4% fewer in the under $30,000 AGI range.

The sample selected under the new design is substantially older than the sample selected under the old design. There are 19% fewer primary filers below age 40 and 23% more primary filers age 50 and older. The number of primary filers age 65 and older is 33% higher and the number of returns with taxable social security income is 43% higher in the new design sample.

6. SUMMARY

Several conclusions emerge from our comparison of the old and new SOI sample designs:

- There are key differences in how the two designs define income. The most important is that, in contrast to the old design, the new design uses net income amounts only when gross amounts are unavailable at the time of selection.

- There are key differences in how the two designs define income strata. The new design adds negative income strata, allowing returns stratified according to negative income to be separated from returns stratified according to positive income. The new design also adds strata for returns judged more useful for policy modeling, allowing such returns to be sampled at higher rates than returns of equal total income judged less useful for policy modeling.

- Compared to the new design, the old design undersamples (1) all negative income returns, (2) high positive income returns, and (3) low positive income returns judged more useful for policy modeling.

- For most income and tax items, the new design supports more precise aggregate estimates, and the gains in precision for some gross income items are substantial.

- For policy modeling, the new design selects a better mix of sampled returns than the old design by selecting more returns with relatively unusual characteristics. The new design provides stronger support for analyzing a wide range of tax policy proposals.

NOTES

[1]The redesign of the cross-sectional sample is part of a larger sample redesign effort that includes the designation of a large panel of tax filing units and the development of annual samples of tax families specified according to the marital and dependency relationships reported on tax returns.

[2]OTA staff provided substantial guidance in defining "useful." Certain returns are deemed more useful because they have relatively unusual characteristics and, thus, their selection would enhance the mix of returns available for policy modeling. A sample consisting entirely of such returns would probably not be useful because it would lack typical returns, the fairly simple returns filed by the vast majority of taxpayers. The critical issue is sample composition.

[3]Tested pension and annuity income equals gross pension and annuity income unless gross pension and annuity income is more than 99 times greater than taxable pension and annuity income, gross pension and annuity income is less than taxable pension and annuity income, or taxable pension and annuity income equals zero. If one of these three conditions is satisfied, tested pension and annuity income equals taxable pension and annuity income. Tested social security income equals gross social security income if gross social security income is less than $75,000 but greater than taxable social security income. Otherwise, tested social security income equals taxable social security income.

[4]Under both designs, business income and expense amounts from up to three Schedule Cs and farm income and expense amounts from up to two Schedule Fs enter the total income definitions separately. Amounts appearing on, for example, a fourth Schedule C are combined with amounts on the third Schedule C.

[5]The new design excludes from total income taxable refunds of state and local income taxes, long-term capital gains from the sale or exchange of a home, and business other income (such as federal and state gasoline or fuel tax credits or refunds).

[6]Only a subset of tax return items are keypunched prior to sample selection, so reliance on net amounts cannot be entirely avoided.

[7]A 1987 sample was not actually drawn using the new design. All new design population and sample estimates for 1987 in this paper are derived from the 1987 SOI sample drawn using the old design.

[8]The 1987 and 1988 business net profit/loss thresholds were $350,000. The net profit/loss amounts from all Schedule Cs filed with a return were combined, and the figure obtained was compared to this threshold. There were 9,593 and 13,852 returns with high business net profit/loss selected in 1987 and 1988, respectively. HINTs are defined by statute. In 1987 and 1988, such returns had adjusted gross income equal to or greater than $200,000, no income tax after credits, and no additional tax for tax preferences. There were 875 and 817 HINTs selected in 1987 and 1988, respectively.

[9]Within each of the 24 income classes, the new design recognizes seven types of returns: (1) Form 2555 attached; (2) no Form 2555 but Form 1116 and either Schedule C or Schedule F attached; (3) Form 1116 attached but no Form 2555, Schedule C, or Schedule F; (4) Schedule C and Schedule F attached but no Form 2555 or Form 1116; (5) Schedule C attached but no Schedule F, Form 2555, or Form 1116; (6) Schedule F attached but no Schedule C, Form 2555, or Form 1116; and (7) no Schedule C, Schedule F, Form 2555, or Form 1116. If a supplementary sample of, for example, returns with foreign earned income (Form 2555) is required, the new design will allow additional returns of that type to be selected.

[10]The reasoning supporting such specifications is discussed at length in Hostetter et al. (1990). One argument is that the richer content of certain returns makes them relevant to a wide variety of potential tax policy proposals. Other returns, in contrast, can be used for simulating only a narrow range of proposals, such as proposals changing the basic tax rates.

[11]To obtain a 1988 new design sample of about 95,000 returns, we adjusted 1987 sampling rates for the strata with 1987 sampling rates under 1% (Strata 7-18), except for Stratum 10 for which the sampling rate of 0.02% was maintained. We required each stratum from 11 to 18 to have the same proportion of sampled returns among those eight strata combined in 1988 as in 1987. (In 1987, 47,370 returns were sampled from Strata 11-18, and 6,393 (13.5%) of those were from Stratum 13. In 1988, 33,928 returns were sampled from Strata 11-18, and 4,571 (13.5%) of those were from Stratum 13.) No sampling rate was allowed to fall below 0.02%. We set the sampling rates for Strata 7, 8, and 9 equal to the sampling rates for Strata 14, 16, and 18, respectively, so that a negative income return would have a sampling rate at least as large as a positive income return with equal income in absolute value.

[12]The 15,000 return difference in sample sizes for 1988 makes comparisons of coefficients of variation difficult. We obtained new design sampling rates for 1988 by adjusting 1987 rates as described earlier. Had we used the 1987 rates for 1988, we would have obtained a 1988 sample of about 108,000 returns. Calculating coefficients of variation for this larger 1988 new design sample, which is still about 2,000 returns smaller than the 1988 old design sample, we find the same patterns among new design and old design coefficients of variation for 1988 as for 1987. For most income and tax items, the new design supports more precise aggregate estimates. Even when the 1988 sampling rates are used and the new design sample is substantially smaller than the old design sample, the new design supports estimates for many income and tax items that are more precise or only slightly less precise than estimates based on the old design. According to Table 2, estimates for several gross income items are substantially more precise, despite the new design's much smaller sample size.

[13]By "simple," we mean a return with income mainly from one source (usually salary and wage income) and no adjustments, itemized deductions, or credits.

REFERENCE

Hostetter, Susan, Czajka, John L., Schirm, Allen L. and O'Conor, Karen (1990), "Choosing the Appropriate Income Classifier for Economic Tax Modeling," in *Proceedings of the Section on Survey Research Methods, American Statistical Association*.

Table 1. New Design Stratum Definitions and 1987 Sample Counts and Sampling Rates

Stratum Definition	Sample Count	Sampling Rate (%)
1. Gross negative income under -9,999,999	583	100.000
2. Gross negative income -9,999,999 under -4,999,999	939	100.000
3. Gross negative income -4,999,999 under -1,999,999	2,040	50.000
4. Gross negative income -1,999,999 under -999,999	1,811	18.500
5. Gross negative income -999,999 under -499,999	863	4.000
6. Gross negative income -499,999 under -249,999	608	1.000
7. Gross negative income -249,999 under -119,999	610	0.400
8. Gross negative income -119,999 under -59,999	492	0.250
9. Gross negative income -59,999 under 0	622	0.100
10. Gross positive income 0 under 30,000 Stratum 10 conditions satisfied.	5,344	0.020
11. Gross positive income 0 under 30,000 Conditions of Strata 10 and 12 not satisfied.	9,681	0.030
12. Gross positive income 0 under 30,000 Useful return conditions satisfied.	8,147	0.080
13. Gross positive income 30,000 under 60,000 Conditions of Stratum 14 not satisfied.	6,393	0.035
14. Gross positive income 30,000 under 60,000 Useful return conditions satisfied.	7,515	0.100
15. Gross positive income 60,000 under 120,000 Conditions of Stratum 16 not satisfied.	4,370	0.080
16. Gross positive income 60,000 under 120,000 Useful return conditions satisfied.	4,258	0.150
17. Gross positive income 120,000 under 250,000 Conditions of Stratum 18 not satisfied.	2,150	0.250
18. Gross positive income 120,000 under 250,000 Useful return conditions satisfied.	4,856	0.400
19. Gross positive income 250,000 under 500,000	5,543	1.000
20. Gross positive income 500,000 under 1,000,000	6,019	4.000
21. Gross positive income 1,000,000 under 2,000,000	8,723	18.500
22. Gross positive income 2,000,000 under 5,000,000	8,123	50.000
23. Gross positive income 5,000,000 under 10,000,000	3,076	100.000
24. Gross positive income 10,000,000 and over	1,319	100.000
25. AGI over $200,000, no income tax after credits, and no additional tax for tax preferences	875	100.000
26. Combined business and farm total receipts $50,000,000 and over	59	100.000

NOTE: Strata 1-9 have negative net income; Strata 10-24 have nonnegative net income. The Stratum 10 conditions are: gross positive income equals net income; both interest and dividend income less than $400; no itemized deductions; no alternative minimum tax; no aged exemptions; filing status not head of household; no dependent children living at home; and no dependent parents. The useful return conditions are: positive alternative minimum tax liability; or gross negative income greater than 40% of gross positive income; or the largest of labor earnings, retirement income, business income, and farm income when combined with interest and dividend income is less than 90% (75% for Strata 16 and 18) of gross positive income. Boundaries for Strata 1-24 are given in dollars.

Table 2. Coefficients of Variation (%) for Selected Income and Tax Components

Income/Tax Component	1987		1988	
	Old Design	New Design	Old Design	New Design
Adjusted Gross Income/Deficit	0.15	0.19	0.15	0.22
Salaries and Wages	0.23	0.28	0.23	0.35
Taxable Interest	0.97	0.98	0.98	1.17
Tax Exempt Interest	5.16	2.11	2.99	2.65
Dividends	1.50	1.37	1.42	1.62
Net Capital Gain	3.77	4.30	3.05	4.27
Net Capital Loss	0.25	0.29	0.28	0.35
Gross Short-Term Capital Gain	3.58	2.06	2.89	2.23
Gross Long-Term Capital Gain	1.12	0.88	0.96	0.98
Gross Short-Term Capital Loss	4.45	1.95	7.55	2.55
Gross Long-Term Capital Loss	5.16	1.66	4.70	2.08
Gross Rent/Royalty Income	2.03	0.95	2.14	1.01
Gross Rent/Royalty Loss	1.94	0.96	2.11	1.10
Gross Partnership/S-Corporation Income	1.54	1.31	1.20	1.06
Gross Partnership/S-Corporation Loss	1.39	1.13	1.30	1.29
Gross Trust/Estate Income	5.94	4.79	5.36	4.77
Gross Trust/Estate Loss	13.20	9.68	12.60	14.39
Taxable Pension/Annuity Income	1.42	1.42	1.44	1.69
Taxable Social Security Income	2.19	1.98	2.23	2.41
Unemployment Compensation	2.27	2.48	2.53	3.15
Net Other (Form 4797) Gain	4.59	3.76	4.10	3.73
Net Other (Form 4797) Loss	6.31	4.90	6.26	6.73
Positive Net Other Income	3.11	2.52	2.60	2.36
Negative Net Other Income	2.31	1.68	2.04	1.68
Total Statutory Adjustments	1.37	1.45	1.42	1.85
Total Itemized Deductions	0.46	0.49	0.50	0.64
Total Tax Liability	0.23	0.27	0.20	0.30
Tax Preference Items	4.41	3.45	3.09	2.43
Income	0.27	0.14	0.26	0.16
Gross Positive Income	0.21	0.14	0.20	0.15
Gross Negative Income	1.03	0.41	1.09	0.62
Combined Total Receipts	1.54	1.29	1.55	1.81
Net Farm Profit	4.64	3.73	4.66	4.86
Net Farm Loss	3.22	2.97	3.35	3.51
Net Business Profit/Loss	1.47	1.43	1.42	1.71
Business Total Receipts	1.66	1.17	1.66	1.80
Depreciation Deduction	2.17	1.21	2.61	1.80
Interest Expense	2.80	1.63	2.95	2.38
Employee Benefit Programs Expense	7.55	4.79	7.86	7.03
Business Gross Profit/Loss	1.35	0.67	1.34	0.97
Sample Size (Number of Returns)	93,343	95,019	110,361	94,973

NOTE: Coefficients of variation for the last six items are estimated from returns with Schedule Cs. There are 30,530 and 31,454 such returns under the old and new designs for 1987 and 37,285 and 32,025 such returns under the old and new designs for 1988. "Income" equals gross positive income if net income is nonnegative and gross negative income if net income is negative. The old design sampling rates used throughout this paper for 1987 pertain to a base year panel sample designated in 1987 and drawn from the full cross-sectional sample of nearly 126,000 returns.

Table 3. 1987 Sample Counts

	Design	<0	[0,15)	[15,30)	[30,60)	[60,120)	[120,250)	250+	Total
					AGI ($1,000)				
overall	old	5574	19068	12704	16246	8910	5706	25136	93345
	new	7059	17024	12794	15495	9872	6897	25877	95017
primary filer's age									
0-22	old	31	6278	470	53	8	13	52	6906
	new	63	4222	325	50	10	15	56	4741
23-29	old	106	3637	2805	1759	336	80	365	9088
	new	149	2624	2056	1344	332	68	316	6889
30-39	old	725	2583	3412	5023	2173	970	3657	18543
	new	923	2472	3049	4112	2033	1006	2695	16290
40-49	old	1773	1525	2161	4331	2818	1834	7209	21651
	new	2180	1715	2254	3934	2741	2140	6648	21612
50-54	old	829	608	824	1487	1051	708	3248	8755
	new	1093	715	1058	1565	1128	766	3413	9737
55-59	old	773	604	785	1336	925	676	3244	8343
	new	937	707	1002	1466	1342	811	3507	9771
60-64	old	600	792	787	1028	647	518	2769	7140
	new	801	990	1020	1326	755	792	3137	8820
65-69	old	379	981	577	515	424	355	1870	5100
	new	449	1247	852	657	779	484	2449	6916
70+	old	347	1877	863	696	520	547	2681	7531
	new	434	2148	1158	1024	745	812	3615	9936
dependent children at home	old	2589	3787	4768	8703	4929	3028	12761	40565
	new	3304	3977	5071	7737	4914	3387	11837	40227
filing status									
single	old	895	12568	4752	2315	722	633	2876	24761
	new	1226	9715	3910	2366	903	847	3050	22018
married filing jointly	old	4337	3849	6323	13187	7986	4906	21193	61783
	new	5481	4915	7426	12407	8742	5834	21817	66622
married filing separately	old	230	373	213	114	40	61	497	1527
	new	229	380	176	118	38	55	492	1488
head of household	old	106	2265	1409	624	161	103	550	5218
	new	116	1997	1271	595	188	159	494	4820
widow(er)	old	6	14	6	7	*	*	20	56
	new	7	17	10	8	0	3	25	69
tax exempt interest	old	728	156	273	804	1250	1589	11087	15887
	new	872	214	415	1129	1791	2028	13156	19605
dividends	old	3279	2101	2410	5642	5223	4325	21304	44285
	new	3877	2810	3152	6093	5956	5165	22911	49964
taxable pension/annuity income	old	584	2247	2045	3134	1929	1023	3644	14606
	new	677	2763	2904	3893	2258	1325	4517	18337
taxable social security	old	21	10	315	1370	895	796	3319	6726
	new	32	15	478	1983	1506	1115	4506	9635
unemployment compensation	old	34	1094	1118	1213	275	43	62	3838
	new	43	1059	1061	1003	248	69	59	3541
statutory adjustments	old	910	837	1851	3393	2313	2081	10292	21677
	new	1115	1236	2545	4043	3208	2537	8072	22756
tax credits	old	8	733	1410	2570	1536	1062	6775	14095
	new	7	762	1627	2494	1790	1560	8829	17069
child care credit	old	0	539	1235	2355	1128	366	952	6576
	new	0	458	1099	1839	982	420	641	5438
foreign tax credit	old	*	*	31	71	242	371	2885	3619
	new	3	37	80	174	338	618	5009	6260
tax preference items	old	1137	38	67	389	956	1490	11665	15741
	new	1389	79	104	755	1594	1947	13141	19010
alternative minimum tax	old	734	19	27	101	246	466	2085	3678
	new	857	24	50	294	404	724	3011	5363
itemized deductions	old	0	1293	4245	12280	8347	5504	24747	56415
	new	0	1938	5075	11453	9141	6545	25508	59660
gross short-term capital gains	old	1622	166	266	817	1155	1518	10812	16356
	new	1870	434	469	1179	1800	1911	12662	20326
gross short-term capital losses	old	1826	201	294	843	1234	1671	11712	17781
	new	2282	406	524	1275	1923	2164	13666	22239
gross long-term capital gains	old	3947	1187	1401	3257	3828	3910	20567	38098
	new	4690	2046	2387	4306	4918	4849	22762	45959
gross long-term capital losses	old	2324	376	494	1289	1738	2078	12976	21275
	new	2942	825	912	1903	2587	2591	15060	26818
gross rent/royalty income	old	3132	915	1209	2588	2616	2704	14054	27218
	new	4074	1806	2176	3781	3349	3716	15423	34325
gross rent/royalty loss	old	2939	844	1141	2504	2527	2631	13449	26035
	new	3796	1660	2018	3682	3215	3570	14722	32663
gross partnership/S-corporation income	old	3154	270	349	925	1518	2498	16103	24816
	new	3525	626	594	1422	2194	2981	17901	29243
gross partnership/S-corporation loss	old	3991	270	405	1129	1973	2979	18336	29083
	new	4457	522	644	1672	2460	3384	19781	32921
business total receipts	old	3150	2286	2256	4019	3077	2255	12234	29276
	new	3726	3254	3178	4528	3727	2981	8978	30373

Note: A "*" indicates that the cell value was suppressed to avoid disclosure of information for specific taxpayers.