Statistics of Income Partnership Studies: Evaluation of the Expanded Sampling Plan

Paul McMahon, Internal Revenue Service

The Internal Revenue Service's Statistics of Income partnership studies are now in their fourth decade. In recent years changes in the tax laws and economy have had a strong effect on this subset of the business population, leading to a need to revise the stratification scheme used in sampling. But these changes also affected the administrative records system we use as the sampling frame, which resulted in still further design adjustments.

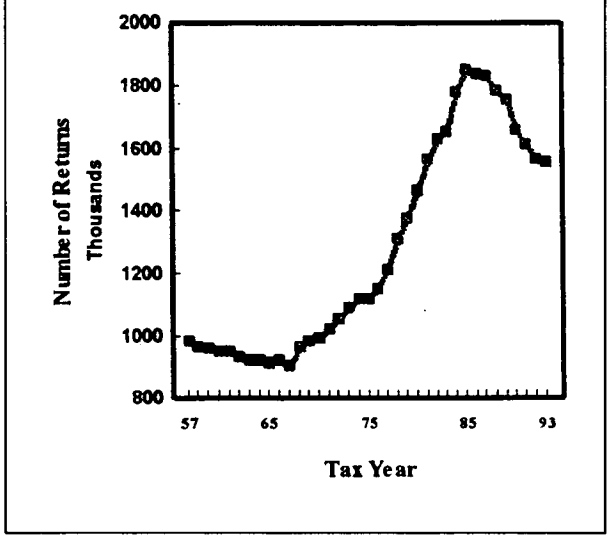
This is the fifth paper in a series documenting changes in the partnership studies. The first three papers (McMahon; Collins and O'Conor, 1990; and McMahon, 1991 and 1993) evaluated the design and subsequent modifications, ending with further proposed changes. The most recent paper (McMahon, 1994) dealt with the creation of a preliminary data file and estimates. Our focus here is on evaluating the effectiveness of the most recent changes, including the effect on the preliminary estimates.

■ Background

The population of interest in these partnership studies is those businesses with more than one owner that, for various reasons, elect not to incorporate. Our interest in this subset of the business community is in its contribution to the overall economic and taxation environment. One of the major clients for these studies is the Bureau of Economic Analysis, of the U. S. Department of Commerce, which uses the summarized statistics in computing the Gross Domestic Product. The other major users, Treasury's Office of Tax Analysis and Congress' Joint Committee on Taxation, use the data files to model the effects of proposed tax bills being considered by the legislature.

Both sets of clients have an interest in the change in the population over the years, so changes in the design risks confounding their analysis of the developing trends. This led us to postpone a major revision in 1987, for example, because we expected the 1986 Tax Reform Act would have quite an impact on the data. As the graph below (Figure 1) shows, it did -- reversing a decade-long run of growth (see also Petska and Nelson, 1990 and Petska, 1991).

Figure 1.--Population Trends



Not all changes could be avoided during the phase-in period of the law. While the overall population size fell, the number of firms in the assets class "$25 million or more" and related high income groups continued to grow. Since companies in these classes were selected with high probability (or certainty), and the budget was relatively fixed, sampling rate adjustments to the smaller strata (as measured by assets or income) reduced the sample in those classes. By 1988 the large classes contained over a third of the sample.

The sample sizes in the smallest asset and income classes had already been reduced as far as we dared, so for the Tax Year 1989 study we cre-

ated an additional stratum. This new stratum capped the older certainty strata at $75 million for total assets and $10 million for income. This allowed us to control the sample size (and cost) through subjecting the records that remained in the original certainty strata to a smaller sampling rate, while maintaining the reliability of the estimates.

But this 1989 change was only meant as an interim response, with a new strata design introduced for Tax Year 1991. This "new" design still used the same stratifying variables and the same sampling frame as the older version. Indeed, to minimize the cost of introducing the strata revisions, we used the same number of strata and their related codes, as in the pre-1989 studies.

As you would surmise, this 1991 revision realigned the strata boundaries, especially the receipts/net income classes. The certainty strata boundary amounts were raised to $100 million for assets and $25 million for income or receipts. We also addressed the limited number of records that were exempt from reporting asset values. As we reported in our 1990 paper, we used a series of regression-based formulae to predict the assets classes for those returns claiming the exemption.

We planned to retain this new version for at least five years, but reliance on administrative systems, such as IRS' Master Files, leaves one at the mercy of the demands on that system. One such demand was to decrease the taxpayer reporting burden on smaller, and especially family-owned, businesses. The regulators decided that the exemption from reporting asset holdings for the smallest firms would be extended. The new rules raised the reporting threshold, relieving more firms of a burden, and dropped constraints (such as requiring that no partner could be a corporation).

Our earlier investigations led us to believe that the asset class prediction method was reasonably stable over the years, but this major liberalization of the exemption caused us to abandon that method. In any case, the prediction formulae proved less stable than we had hoped.

■ The Latest Revision

Most people imagine that every scrap of information the Internal Revenue Service receives is added to vast computer databases. For partnerships, however, only selected data are abstracted from the forms that are filed. A portion of that data is, at best, of slight use for our stratification goals, because an item is either very highly correlated with another variable or arises from purely administrative procedures. Our 1988 design already used most of the information available, and since then several fields have been removed, though none that affected stratification.

Thus, any new design will end up making use of the same items as stratifiers. Moreover, to keep definitions consistent over time, some of the partnership records' items must be combined. One instance of this is the receipts stratifier, which consists of gross receipts plus gross rents (real estate and other) plus portfolio income. Prior to 1987 many of these items were all reported as a single entry on the form, but the law separated some parts of it, because they were considered a result of passive investment (for example, real estate rents, dividends, and long-term capital gains).

The monetary data tend to be of reasonable quality because the structure allows verification through some elementary accounting. The frame's industry codes, on the other hand, which were self reported and, except for the most common categories, often unverifiable, were a continuing source of difficulties. This is why the earlier designs limited the use of the industry code to separating the real estate operators (about a third of the businesses) from other types of firms. Over the past couple of years, though, the IRS has made significant progress in improving the reliability of the industry codes on the files we use as our sampling frame.

We decided to use this improvement in quality to offset the loss of definition on the size of total assets among smaller companies. The design that resulted is shown in Figure 2 on the next page.

Figure 2.--Tax Year 1993 Partnerships, Strata Definitions with Populations

Assets $100,000,000 or more	2,577
Assets less than $100,000,000 and Receipts/Income $25,000,000 or more	2,506

	Real Estate Operators	Farms, Trades, Finance and Services	Mining and all others
Assets $25,000,000 under $100,000,000			
Receipts/Income $5,000,000 under $25,000,000	1,837	1,228	354
Receipts/Income less than $5,000,000	1,706	2,346	212
Assets less than $25,000,000			
Receipts/Income $5,000,000 under $25,000,000	1,576	7,123	1,985

Real Estate Operators

Assets ($)	Absolute Value of Receipts/Income ($)					
	Under 50,000	50,000 under 100,000	100,000 under 250,000	250,000 under 500,000	500,000 under 1,000,000	1,000,000 under 5,000,000
Under 250,000	211,206	37,845	19,769	{ ...	7,003 ... }	
250,000 under 750,000	31,730	33,409	35,425	{ ...	14,400 ... }	
750,000 under 2,500,000	{ ... 17,476 ... }		30,798	25,011	{ ... 18,224 ... }	
2,500,000 under 5,000,000	{ ...	5,263 ... }		6,755	10,103	6,703
5,000,000 under 25,000,000	{ ...	2,516 ... }		1,693	4,649	14,103

Farms, Trades, Finance and Services

Assets ($)	Under 40,000	40,000 under 100,000	100,000 under 250,000	250,000 under 1,000,000	1,000,000 under 2,500,000	2,500,000 under 5,000,000
Under 250,000	424,633	102,039	86,691	71,915	{ ... 13,993 ... }	
250,000 under 750,000	32,258	15,301	11,547	14,553	{ ... 9,009 ... }	
750,000 under 2,500,000	{ ... 22,700 ... }		10,584	11,768	5,225	2,551
2,500,000 under 5,000,000	{ ... 4,267 ... }		2,396	5,095	{ ... 3,962 ... }	
5,000,000 under 10,000,000	{ ...	2,782 ... }		2,766	1,917	1,089
10,000,000 under 25,000,000	{ ...	1,462 ... }		{ ... 2,939 ... }		971

Mining, Construction, Manufacturing and Transportation

Assets ($)	Under 40,000	40,000 under 100,000	100,000 under 250,000	250,000 under 500,000	500,000 under 1,000,000	1,000,000 under 5,000,000
Under 250,000	51,865	24,351	24,559	10,659	{ ... 6,224 ... }	
250,000 under 1,000,000	{ ... 3,276 ... }		1,600	{ ... 3,375 ... }		2,024
1,000,000 under 5,000,000	{ ... 1,095 ... }		576	{ ... 1,599 ... }		2,086
5,000,000 under 25,000,000	{ ...	302 ... }		{ ... 268 ... }		861

There is another major difference between this design's outline and those of past years. The 1992 strata plan used a matrix approach, where a given asset class and a given income/receipts class determined the sample stratum. There are few records with small assets and very large receipts or large assets and small receipts. Under these conditions, Neyman allocation, even ignoring costs, prescribes sample sizes near zero for these marginal strata. This led us to set the probabilities for these classes at the same level as adjacent income classes in the same asset group or at certainty, if the population were small enough.

The 1993 design shows income/receipts class boundaries that depend upon the asset class. This revised strata plan explicitly achieves the same collapsing strategy as the sampling rate adjustments used in 1992.

Since the probabilities of selection vary from under 0.2 percent for the smallest stratum (in monetary terms) to certainty for the largest classes (such as those for firms with assets of $100 million or more), we employ weights that range from over 550 to 1. Because the weights are constant within a stratum, though, they do not increase the variance in measures of level.

But our main clients are interested in change across the years. If the sample is large enough, as at the Industrial Division level, the shifts in weights will not be evident. At a finer level of aggregation, say for a minor industry, a single firm's change in weight might cause a visible deviation. This is not an improbable event, as the sample selection algorithm was designed to increase the likelihood of such occurrences -- see Harte (1986) for a discussion of the process. There are two ways such a change might occur: the firm's economic status may grow (or collapse) or the selection probabilities may change significantly over the years.

The first situation is a problem where the probabilities of selection vary greatly across the strata. The reallocation of sample resources reduced the weights assigned to most classes of records. A key indicator of this is the size of the maximum weight (Figure 3). The second problem -- greatly altered selection probabilities -- was unavoidable under the circumstances. Yet the stepwise reductions we see here ameliorated that effect to some extent.

Figure 3.--Maximum Weights

Tax Year	Real Estate	Other Industries
1989	2,380	1,320
1991	970	800
1992	550	340

■ Results

As we mentioned above, the goal of this revision to the strata design was not to accommodate any new subsidiary study nor to refocus on some subpopulation critical to our clients. Rather, we meant to retain the accuracy of the estimates in the face of a changing regulatory and processing environment. Since the data abstraction process did not change (indeed, we retain most of the same data editors from year to year), we need not be overly concerned about other sources of variation greatly affecting the estimates for the purpose of evaluating the changes.

The coefficients of variation (CV's) in Figure 4 illustrate that the reliability has been preserved for the national, all industries estimates.

Four different strata designs are represented here, 1988, 1989/90, 1991/92 and 1993. The modification for the 1989/90 version is slight, yet the effect is quite pronounced. This "improvement" is actually due to the diffusion of efforts that an overly large "critical case" class inflicted on the 1988 study. Many of these returns were filed very late

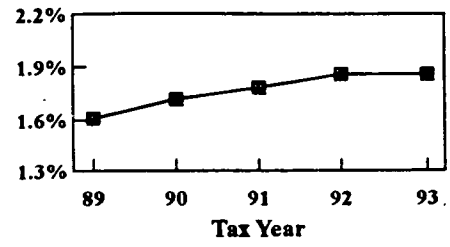
Figure 4.--Coefficients of Variation for Selected Estimates
(percent)

Items	Tax Year					
	88	89	90	91	92	93
Total Assets	1.2	0.6	0.5	0.7	0.4	0.4
Receipts	0.8	0.6	0.4	0.4	0.3	0.2
Net Income	3.0	1.3	1.4	1.8	0.7	0.6
Net Loss	2.2	1.8	1.8	1.0	1.0	1.2
Salaries	....	1.2	2.4	0.8	0.7	0.6
Depreciation	1.4	1.2	1.2	0.8	0.7	0.8

in the year, so there was little time to search for them before the study ended. As a result, about one percent went unlocated before the data abstraction cut-off. That is, the coefficients of variation for 1988 include a sizeable adjustment for a nonresponse problem that was all but eliminated in later years.

There is a series of marginal improvements in these estimates after 1988. This might be the result of a significant increase in the ratio of total sample to total population, and it is clear that the population is decreasing. But the sample size has been roughly constant and the population hasn't shrunk that much. Figure 5 confirms that although there may be some small increase in that ratio, it is far from demonstrating that this is the source of the estimates improved reliability.

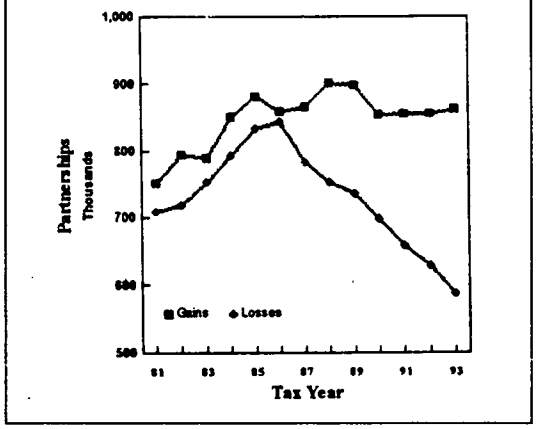
Figure 5.--Sampling Fractions



The first pair of items in Figure 4, assets and receipts, are stratifiers so we would expect reasonable performance by the design in these areas. In fact, this is what we see, although the improved coefficients for receipts are better than expected.

The trends for estimates of net income and net loss are of real interest, and so are separated here. Over the six years, the CV's for both have generally improved. For net income, this may be due to an increase in the number of sample records showing a profit, which is reflected in the population estimates of Figure 6. (Indeed, for the first time in decades, the 1993 estimated amount of net income exceeds net losses.) Given this trend, and an overall sample size that is constant at about 29,000 active firms, the converse is also true, that the sample of firms showing losses is decreasing. Thus, the surprising finding is that the estimates for net loss have improved through 1991 and are stable since then.

Figure 6.--Gains Versus Losses



The other set of estimates in Figure 4, salaries and depreciation, are not used in the stratification and are not overly correlated with those items. We use them as proxies to evaluate the performance of the preponderance of estimates. Here we see improvement, which we believe relates to the increased allocation of sample to the smaller strata, which resulted in the shrinkage of the range of the weights.

Another view of the quality of the estimates comes from the effect of the designs on the industrial division data. The sample allocation for the 1990 study was strained by the changes in the population that had arisen since its strata plan was first employed. One result was that we were unable to focus resources on the industry divisions. The 1990/91 studies focused attention on these domains through adjustments to the sample allocation in smaller firms' strata, while the 1993 study, as we have seen, provides separate strata for the smaller divisions.

Figure 7.--Coefficients of Variation by Industrial Division
(percent)

Industry Division	1990	1991	1992	1993
	Total Assets			
All	0.5	0.7	0.4	0.4
Agriculture	8.1	5.2	4.1	3.7
Mining	9.9	2.2	2.0	1.3
Construction	9.1	5.4	4.5	2.9
Manufacturing	10.0	1.0	0.7	0.6
Transportation	10.9	1.5	1.3	0.9
Trade	5.7	1.7	1.7	1.8
Finance etc.*	1.2	1.7	0.6	0.6
Real Estate Op's	0.9	0.9	0.9	0.9
Services	1.4	1.1	1.0	1.1
	Salaries and Wages			
All	2.4	0.8	0.7	0.6
Agriculture	23.1	9.2	9.1	8.4
Mining	11.1	7.4	7.3	6.6
Construction	11.2	7.4	7.4	5.5
Manufacturing	2.3	1.9	1.5	1.5
Transportation	4.3	2.6	2.2	2.0
Trade	3.2	2.3	2.3	1.9
Finance etc.*	23.9	2.8	2.4	2.4
Real Estate Op's	6.9	8.2	10.4	10.5
Services	1.3	1.1	1.0	0.9

* Excludes Real Estate Operators

The assets data in Figure 7 show that the adjusted allocation for the 1991 and 1992 studies did improve the divisional estimates. The change from an implicit focus to a separate strata, in 1993, for the smaller divisions further improved those estimates, reducing the CV's by about a quarter, while having little discernable effect on the other divisions.

The salaries data in that table also show significant improvement over the 1990 results for most divisions. The difference between the 1992 and 1993 results are not as strong as for the assets data. In part this is because labor costs data can be reported in many places on the partnership tax return, and here we are inspecting only one. In the case of real estate operators, for example, the information is frequently reported as a deduction on the rental schedule, so the estimated amount is rather small considering the size of that industry.

In this analysis we have treated real estate operators as if it were a division. The population of partnerships in this industry is greater than that of the rest of finance, insurance and real estate or any of the other divisions, which is reason enough. This separate treatment, though, allows us to focus better on the effects of the design changes, for the separate real estate strata have a twenty year history.

■ Effect on Preliminary Estimates

As we mentioned above, there are two main products for the partnership project. So far we have discussed the data from the full sample, which is published in the fall issue of the *Statistics of Income Bulletin* each year (see, for example, Wheeler, 1994). The other product, preliminary estimates, is not available for public use at this time, both because the methods are experimental and the source

data have not been thoroughly verified. Outlier detection processing, for example, is only applied to the final data file, both for timeliness and because records that would seem to be outliers early, look quite ordinary later, due to the pattern of large firms filing at the end of the year.

Figure 8.--Error in Preliminary Estimates

Items	Relative Error, in Percent			1993
	1991	1992	1993	CV
Partnerships	0.8	2.0	0.2	0.3
No. of Partners	0.7	7.7	1.0	2.7
Total Assets	-4.4	2.2	0.5	0.4
Receipts	-4.7	1.0	3.9	0.2
Net Income	-0.6	4.8	-1.6	0.6
Net Loss	-12.3	-10.1	-13.7	1.2
Portfolio Inc.	5.4	1.6	2.2	...
Depreciation	-11.8	-1.8	1.3	0.8
Taxes Paid	-6.5	-0.7	4.7	0.7
Sal. & Wages	1.5	6.3	1.0	0.6

The relative error described in Figure 8, above, is the difference between the preliminary and final estimates divided by the latter. We include the coefficients of variation for the 1993 final estimates as a standard against which the preliminary estimates may be measured.

There are two sources of error for these preliminary estimates, as we described in last year's paper (McMahon, 1994). The first source is from having to predict the population sizes and the second from the early sampling cut-off. Predicting the total population size and distribution with respect to assets seems to be well in hand. The receipts/income classes, however are another matter, as illustrated in the 1993 data, where we overestimated the number of firms in the larger monetary classes. This led to the overestimate of Receipts by 3.9 percent. The taxes paid deduction, which closely parallels the receipts figures, was also overstated, by 4.7 percent, though no causal link is immediately apparent.

The largest problem of all, though, is in estimating net loss, which has a relative error that continues to run at least ten percent below the final figure. The understatement is due to the propensity for the firms with large losses to file late in the filing period. While returns are, in most cases, due April 15, an automatic six month extension is granted upon request (because there is no tax liability, especially when there is a loss to be reported). Once an allowance is made for mail delivery and the administrative data abstraction process, we see that few returns with extended due dates arrive in time for our study's early data collection cut-off.

■ Further Research

The understatement of net losses in the preliminary statistics is a measure of how nonsampling errors can affect our estimates. In the case at hand, the problem was self-inflicted nonresponse caused by the early cut-off. But there are other non sampling problems present in the system. A taxpayer may report, for example, a farrier's fees in "other deductions." In our data abstraction and editing process it should be reassessed as a farm expense, but perhaps the editor was raised in Manhattan, and deleted the data as an unallowable business expense for a fur coat (mistaking farrier as furrier) ... this case was corrected by an alert senior editor.

We have systems in place to detect and prevent such problems, but have not reviewed the effect on the estimates in several years.

The strata plan now in use depends on the population distribution across the current industry codes. These industry codes, however, are being replaced with an entirely new set of codes with a very dif-

ferent underlying philosophy. How will these new codes affect the sample design? The implementation date is some years off, but given the project's lead time requirements, we should begin studying the expected affect in the next year.

■ References

Harte, James M. (1986). Some Mathematical and Statistical Aspects of the Transformed Taxpayer Identification Number: A Sample Selection Tool Used at IRS, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

McMahon, Paul (1994). Statistics of Income Partnership Studies: Evaluation of Preliminary Estimates, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

McMahon, Paul (1993). Statistics of Income Partnership Studies: Evaluation of the Revised Sampling Plan, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

McMahon, Paul (1991). Statistics of Income Partnership Studies: Sampling Plan Redesign II, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

McMahon, Paul; Collins, Richard.; and O'Conor, Karen (1990). Revising the Statistics of Income Partnership Sampling Plan, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Petska, Tom (1991). Further Examination of Tax Shelters in the Post-Reform World, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Petska, Tom and Nelson, Susan (1990). Partnerships and Tax Shelters: An Analysis of the Impact of the 1986 Tax Reform, *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Wheeler, Tim (1994). Partnership Returns, 1992, *Statistics of Income Bulletin*, Fall 1994, Internal Revenue Service. ■