THE 1995 TAXPAYER COMPLIANCE MEASUREMENT PROGRAM (TCMP) SAMPLE REDESIGN - A CASE HISTORY

Wendy Rotz, Jeri Mulrow, Eric Falk
Wendy Rotz, Internal Revenue Service CP:R:S:P:S, P.O. Box 2608, Washington D.C. 20013-2608

KEY WORDS: Stratified Sampling, Optimum Allocation, Industry Codes

Introduction

The Taxpayer Compliance Measurement Program (TCMP) is periodically conducted by the Examination Branch of the Internal Revenue Service (IRS) to estimate compliance with tax laws and revenue lost from noncompliance. Random in-depth audits with intensive probes for underreported income are performed. The resulting data are used for subsequent audit selection strategies and to study proposed tax law changes. The largest TCMP study in the history of the IRS will begin in 1995 and will have significant changes from historical TCMP studies: industry will now be a primary stratifier; all major income tax forms will be audited the same year; and there is particular interest in studying compliance by geographical area.

Many difficulties arose while incorporating the changes. There has never been a uniform definition of compliance that applies to all of the major tax forms. Each form differs in its detail and scope of industry codes. Taxpayers frequently misclassify themselves or fail to report their industry at all. Some industries are too sparse to be studied in small geographic areas and only obsolete audit data were available for portions of the redesign.

This paper discusses practical solutions to these problems, summarizes the final design, and outlines possible approaches to future TCMP studies.

Background

The first TCMP study began in tax year 1963 with an extensive probe of 92,000 individual Form 1040 tax returns. Since then, there have been almost fifty TCMP studies of various tax returns and special issues. Studies of major forms and issues have been repeated every few years. However, there has not been a TCMP study since tax year 1988 when individual Form 1040 returns were sampled. (See Figure 1.)

Within the IRS, the primary use of TCMP data is in developing audit selection strategies to identify groups of returns with a high probability of more taxes owed. The data are also used to launch taxpayer education programs

Figure 1: Historical TCMP Studies

Form Type	Tax Year	Sample Size
1040 Individual	1963	92,000
	1965	50,000
	1969	53,000
	1971	26,000
	1973	55,000
	1976	50,000
	1976	55,000
	1979	50,000
	1985	50,000
	1988	54,000

Form Type	Processing Year	Sample Size
1120 Corporations	1969	16,000
	1973	20,000
	1978	33,000
	1981	33,000
	1988	19,000
1120S Corporations	1985	10,000
1065 Partnerships	1982	27,000

and to estimate voluntary compliance levels, defined as the ratio of tax reported to the sum of tax reported and the after audit tax increase. Other government agencies that use TCMP data include the Office of Tax Analysis (OTA) in the Department of Treasury, the Bureau of Economic Analysis (BEA) in the Department of Commerce, the Joint Committee on Taxation (JCT) in Congress, and the General Accounting Office (GAO). They have used TCMP data to evaluate enforcement, consider tax law changes and to estimate the national tax gap, defined as the difference between the revenue actually collected by the IRS and the revenue that would be collected if taxpayers reported all taxes owed.

Historically, business TCMP samples have had twelve to fifteen size strata, categorized by variables like gross receipts, total assets, and various versions of positive income. Huge companies may operate at a net loss; therefore, these stratifiers were chosen to indicate the size of a company, regardless of its taxable income. However, individual 1040 TCMP studies from the 1980s suggested industry and geographic location may have an effect on noncompliance. [2,5,6] These findings were the impetus to redesign the sample and stratify future studies based on industry and location. In addition, this choice of stratification corresponded with an IRS initiative for auditors to become experts in their local industries.

In previous years, TCMP studies focused on one tax form at a time. However, it is now desired to study compliance issues for an entire industry, regardless of how companies organize and pay their taxes. Therefore, in 1995, all major tax forms will be studied concurrently. They include: the individual 1040 family (1040, 1040A, 1040PC and 1040EZ); corporate Forms 1120 and 1120A; the partnership Form 1065; and the S-corporation 1120S Form. In addition, Foreign Controlled Corporations (FCCs) will now be sampled as a separate stratum. In the past, FCCs were sampled together with the other 1120/1120A returns.

The Statistics of Income (SOI) began working on the TCMP sample redesign in the summer of 1993 at the request of the IRS Research Division [3], which had traditionally designed the TCMP studies. Throughout the consulting process, Research made the final decisions. Other customers were allowed to have some input, however, SOI's communication was restricted to internal IRS employees. Therefore, SOI had to rely upon Examination and Research to address all user needs.

A working group with representatives from Research, Examination, and SOI was formed to discuss user needs, which were often conflicting, and design feasibility. The group met weekly and was critical to the redesign. Without the cooperative effort, the time frame could not have been met and the conflicting needs would not have been resolved. In addition, an outside independent contractor, Klemm Analysis Group, reviewed the statistical approach and final designs. [4]

Original Specifications of the New Design

By the time SOI was consulted, compliance researchers had already categorized the nation's industries into thirty strata, such as farming, manufacturing, and wholesale trade. They had also divided the country into thirty geographic areas. This created 900 strata and, therefore, SOI was requested to reduce the twelve to fifteen historical size categories to three, making a total of 2,700 strata.

In addition, all major tax forms were to be sampled and compliance data combined across the forms. The sample was to provide a minimum number of returns with a high after audit increase in total tax . Sample weights were not to exceed a ratio of 1:5 within certain groups of strata. Moreover, the sample was to remain comparable to past TCMP studies.

A Uniform Definition of Compliance

The first obstacle encountered has been a flaw since the project's conception:

> To date, there does not exist a compliance measurement that is meaningful across all tax forms.

This is due to the inherent nature of the forms. Taxes are paid on 1040 and 1120 returns. By contrast, the 1065 and 1120S are primarily information returns; partnership and S-corporation taxes are usually paid on the individual returns of the partners and shareholders. This fundamental difference causes few variables to be present across all form types. Those that do appear may vary in definition; for example, a variable as basic as "income" is defined differently for Forms 1120 and 1065. Taxable income was briefly considered as a component of a new compliance measure because it is similarly defined and appears on all forms. However, it was quickly rejected because it is highly volatile and a poor measure of compliance. A consistent, meaningful, universal definition of compliance has been elusive. Ultimately, perhaps only data from similar forms will be combined.

The absence of a compliance definition and operational considerations led to separate designs for each type of return. In all, six different sample designs were developed for the 1995 TCMP study, one for each distinct type of return: individual non-business 1040 returns, individual business 1040 returns, corporate 1120/1120A returns, FCC returns, S-corporation 1120S returns, and partnership 1065 returns.

Consistent Categorization of Industries

Another problem combining across forms is consistent placement of the Principal Business Activity (PBA) codes that are used to identify type of industry. Each of the Forms 1040, 1120 and 1065 has a unique set of industry codes. The 1040 forms are heavy with service codes, while the corporate 1120 forms have extensive detail in manufacturing. Occasionally, it was difficult to place the same industry in the same stratum across all form types. For example, on Forms 1040 and 1065 veterinary services are under agricultural services; however, veterinarians cannot even be identified on Form 1120 because they fall under miscellaneous services.

Where possible, the Standard Industrial Classification (SIC) was used as a guideline for consistent placement of the industry codes and some of the original thirty industry strata were consolidated. Despite these efforts, there are

still a few similar industries split between different strata across the form types.

Misclassified Industries

Taxpayers are not required by law to report their industry codes. The raw self-reported codes are frequently missing, invalid, or misclassified creating extensive reclassification after audit. For example, on 1040 returns, only 83 percent of the taxpayers indicating a heavy construction code actually belonged in this stratum, 14 percent were reclassified as building trade contractors and three percent were reassigned to a variety of other strata. Furthermore, based on raw industry codes, almost twenty percent of the business 1040 returns fell into a catch-all category for missing, invalid or unable-to-classify codes. Yet, after editing these codes, the percentage fell to less than one percent.

It was not feasible to pre-screen industry codes prior to sampling since there are millions of returns filed each year, although, the IRS is making some effort to improve the accuracy of these codes. Even if edited codes were available prior to sampling, it may not be desirable to sample TCMP returns by them. A fundamental objective of TCMP is to compare audited returns to raw taxpayer returns. Any editing prior to audit would defeat this purpose.

The misclassification problem was partially reduced by further collapsing some of the proposed industrial groups into broader strata. Sample sizes were adjusted and inflated to account for a certain proportion of returns being reassigned to other strata. However, this approach was costly, not just from the additional sample, but also from the resources drained by strata for miscellaneous codes and industries. For example, the catch-all category mentioned above consumed twenty percent of the 1040 business resources during allocation.

After sampling, misclassified returns will cause weighting issues during analysis. Raking may be needed to adjust weights when analyses are desired for reclassified industries. However, future audit selection formulas will most likely be based on raw industry codes and will, therefore, have less weighting considerations.

Historical Data

The most recent 1120/1120A and 1040 TCMP studies were conducted for tax years 1987 and 1988 respectively. The minor tax law changes since then were compensated for during the sample design. However, the only 1065 and 1120S TCMP studies were conducted in 1982 and 1985 prior to the Tax Reform Act of 1986. It was felt that ten plus years and major tax law changes made these data obsolete for the purpose of designing a stratified sample based on optimum allocation. (Optimum allocation requires a reasonably accurate knowledge of standard errors and population characteristics [1].)

However, current population estimates were available from SOI for tax year 1991. Corporate, individual, and partnership data are collected annually by SOI, primarily for economic analysis. Since these files do not include audit data, SOI records could only be used for population estimates.

Therefore, sample designs for S-corporations and partnerships were developed without the benefit of variance estimates. Proportional allocation was suggested for easy implementation. However, Research chose equal allocation to ensure a minimum sample size in each industry stratum. Based on resources and user interests, Research also specified a total sample size of 12,500 returns for each of these two forms.

Difference in total tax, defined as after audit total tax minus before audit total tax, is a major component of compliance and became the key variable of interest used in the sample allocation of 1040 and 1120 returns. The sample design for FCC returns simply used total tax for sample allocation because FCC audit data for the difference in total tax were not available.

Size and Location Stratifiers

Initially, requirements of the new design specified the primary stratifiers to be industry group and geographic location. However, certain industries are sparse in particular geographical areas. For example, farms are rare in Manhattan. Reliable sampling rates by location would have been difficult to determine with available data. Furthermore, the historical size stratifiers produced better variances and separated returns with large after audit tax increases better than the newly proposed industry strata. Moreover, historical comparisons would be extremely difficult without historical size strata in the sample design.

Therefore, a compromise was negotiated. A national sample would be designed. and stratified by industry and historical size stratifiers. It would be poststratified by location. Final estimates will only be provided for industries prevalent in a given area. Analysis of sparse industries will require a larger geographic area.

The Final Design

To summarize the developments:

- The 1995 TCMP study will consist of six designs, one tailored for each type of return: individuals without businesses, individuals who own their own business, corporations, FCCs, S-corporations, and partnerships.

- New industry strata have been incorporated into the design, but all historical size stratifiers have been preserved.

- The data will be poststratified by geographic location analyzing sparse industries over broad areas.

- Most industry strata will be oversampled to compensate for classification errors.

In addition, adjustments were made to the sample after allocation to obtain a minimum sample per substratum and a minimum number of high change returns. Due to the increased number of strata and resource constraints, only 1040 non-business sample sizes were increased to achieve a satisfactory ratio of weights.

Individual and corporate sample sizes were determined using optimum allocation. Partnership and S-corporation sample sizes were equally allocated. Stratification of the 1040 non-business portion of the TCMP sample remained similar in structure to previous TCMP studies.

Refer to Figure 2 for a summary of the final 1995 TCMP sample design. A total of 153,900 returns will be sampled in 789 strata. While this is highly stratified, it is a significant reduction from the originally proposed 2,700 strata. The different coefficients of variation were based on resources and user interests. The only available partnership and S-corporation data were obsolete and could not be used for variance estimates.

Next Steps

There are still many questions that will need to be resolved before and during analysis of the 1995 TCMP data: 1) What is a meaningful measure of compliance across form types? 2) Are reweighting methods sufficient to account for industry misclassifications? 3) How should data be partitioned by geographical location or consolidated by industry for analysis? Many more questions are sure to arise as the data are collected and analyzed.

There are several directions that the IRS could take with future TCMP studies. All have advantages and disadvantages. One approach would be for the IRS to continue to perform large scale TCMP studies on a periodic basis. Historically, it takes over three years to design, capture data, and perform analyses for a TCMP study. The IRS could complete a TCMP study before embarking on a new one. Lessons learned from complete data analyses could be incorporated into future studies and resources could be allocated one study at a time.

Figure 2: 1995 TCMP Sample Design

Form Type	Size Stratifiers	No. of Industry Strata	Total No. of Strata	Expected CV	Approx. Sample Size	Approx. Pop. Size	% of Pop. in Sample
1040 Non-business	Total Positive Income Taxable Income Exemptions Non-business Total Positive Income	0	12	3.5%	34,000	111,000,000	0.03%
1040 Business	Total Gross Receipts Schedule C and F Non-business Total Positive Income	23	345	1.5%	58,900	7,000,000	0.84%
1120 / 1120A	Total Assets Gross Receipts	21	294	2.0%	35,300	2,300,000	1.53%
1120S	Total Assets	21	42	N/A	12,500	1,800,000	0.69%
1065	Gross Receipts No. of Partners	22	66	N/A	12,500	1,500,000	0.83%
1120FCC	Total Assets Gross Receipts	5	30	12.5%	700	57,000	1.23%
TOTALS			789		153,900	123,657,000	0.12%

A disadvantage of this approach is the absence of randomly selected compliance data in years without a TCMP study. It is difficult to study trends without annual data. Also, tax law changes may dramatically alter population characteristics between studies making historical comparisons difficult.

An alternative could be to use the 1995 study as a baseline and augment the data annually with much smaller studies. This approach would give the IRS enough data to monitor trends year after year, and to target certain areas for further study. Smaller, more flexible studies would be less of a resource burden and could still provide interesting data. Unfortunately, smaller annual studies may not provide enough data to update audit selection formulas or to produce estimates for all industries, particularly after poststratification by geographic area.

Yet another option may be to integrate the regular annual audit data with periodic TCMP data. However, regular audits are not randomly selected. There is a statistical problem in determining how or even if random and non-random data can be combined to produce desired estimates. A similar strategy could be to augment TCMP data with non-audit data that is regularly collected for other purposes. Again, the feasibility needs to be explored.

Now is the time to begin designing for the future. The 1995 TCMP sample redesign has been completed. The selection of returns for the 1995 study will begin in January 1995 and run through December 1995. Audits will begin sometime after July 1995 and will take two years to complete. Preliminary analysis and results are not anticipated until 1997 or 1998. By that time, several small TCMP studies could be developed or the planning of the next large study should already be underway.

References

[1] Cochran, William G. (1977). *Sampling Techniques*. New York: John Wiley and Sons, Inc.

[2] Christian, Charles (1992). "Compliance of Sole Proprietors--Findings from 1988 TCMP Phase III, Cycle 10," *Publication 1500*, pp. 23-28.

[3] General Accounting Office (1993). *Tax Administration: IRS' Plans to Measure Tax Compliance Can Be Improved*, GAO/GGD-93-52.

[4] Klemm Analysis Group (1994). *Review of Sample Design for Taxpayer Compliance Measurement Program Final Report*, Unpublished.

[5] Koteen, Gloria (1986). "Regional Trends in Individual Voluntary Compliance Levels for 1965-1982," *Trend Analyses and Related Statistics*, Doc 6011 Washington D.C. IRS, pp. 27-34.

[6] Plumley, Alan (1986). "Trends in Industries With Low Compliance in Reporting Business Receipts," *Trend Analyses and Related Statistics*, Doc. 6011 Washington D.C. IRS, pp. 111-115.