

June 2001

TAX  
ADMINISTRATION

Status of IRS' Efforts  
to Develop Measures  
of Voluntary  
Compliance



G A O

Accountability \* Integrity \* Reliability

---

# Contents

---

<b>Letter</b>		1
	Results in Brief	3
	Background	4
	IRS Has Not Finalized Its Plans on How It Will Measure Voluntary Reporting Compliance	5
	How Other Programs Measure Compliance	9
	Agency Comments	11
<b>Appendix I</b>	<b>Objectives, Scope, and Methodology</b>	13
<b>Appendix II</b>	<b>TCMP Studies</b>	15
<b>Appendix III</b>	<b>Summary of GAO Assignments Related to Measuring Voluntary Compliance</b>	16
<b>Appendix IV</b>	<b>How Compliance Is Measured for Six Programs</b>	17
<b>Appendix V</b>	<b>Comments From the Internal Revenue Service</b>	24
<b>Tables</b>		
	Table 1: Objectives and Guiding Principles for the NRP Program	6
	Table 2: Compliance Review Characteristics for Six Federal Programs	10
	Table 3: Agencies Contacted and Programs Administered	14
	Table 4: Types of TCMP Surveys by Return Type and Tax Year	15

---

---

## **Abbreviations**

CERT	Comprehensive Error Rate Testing
CMP	Compliance Measurement Program
HCFA	Health Care Financing Administration
GPRA	Government Performance and Results Act
INS	Immigration and Naturalization Service
IRS	Internal Revenue Service
INTEX	Inspections' Traveler Examinations
NRP	National Research Program
OASDI	Old Age, Survivors, and Disability Insurance
SSA	Social Security Administration
SSI	Supplemental Security Income
TCMP	Taxpayer Compliance Measurement Program



United States General Accounting Office  
Washington, DC 20548

June 18, 2001

The Honorable Max Baucus, Chairman  
The Honorable Charles E. Grassley  
Ranking Minority Member  
Committee on Finance  
United States Senate

The Honorable William M. Thomas, Chairman  
The Honorable Charles B. Rangel  
Ranking Minority Member  
Committee on Ways and Means  
House of Representatives

The U.S. tax system is based on voluntary reporting—taxpayers self-report their liabilities on the returns they file. The Internal Revenue Service (IRS) reviews all tax returns after they are filed to ensure compliance with the tax laws governing this voluntary system. These reviews include the math error program, matching of information returns to tax returns and operational audits. Operational audits are targeted on returns where IRS suspects noncompliance, often based on a computerized screening of returns that looks for indicators of noncompliance.<sup>1</sup>

In spite of IRS' efforts to enforce compliance, each year billions of dollars in taxes owed are not voluntarily reported and paid, which could result in reduced revenue to fund federal programs, higher tax rates, or both. In order to understand the overall extent of noncompliance and the effectiveness of its compliance assurance programs, IRS has periodically measured voluntary compliance. There are three types of voluntary compliance measures: filing compliance, which measures the percent of taxpayers who file returns in a timely manner; payment compliance, which measures the percent of tax payments that are paid in a timely manner; and reporting compliance, which measures the percent of actual tax liability that is reported accurately on returns. This report reviews reporting compliance.

---

<sup>1</sup>Math error checks are electronic reviews of a tax return that identify clerical or mathematical errors, including incorrect Social Security numbers. Information return matching compares information on income provided by payers with information reported by taxpayers on their return. Audits are examinations of taxpayers' returns in which IRS requests additional documentation to verify what is on the return.

---

IRS' past estimates of reporting compliance were based on audits conducted on random samples of taxpayers that were then statistically projected to the entire population of taxpayers. Unlike operational audits, done to help ensure compliance and targeted on suspected noncompliance, random audits are done to measure compliance and are conducted on a sample of returns selected without regard to suspected noncompliance.

IRS does not have a current measure of voluntary reporting compliance. The last time IRS measured voluntary reporting compliance was for tax year 1988 returns. IRS planned to measure voluntary reporting compliance for tax year 1994, but canceled the effort because of criticism that the project was too costly and random audits were intrusive and burdensome on compliant taxpayers. However, the Commissioner of Internal Revenue has said that a current measure of voluntary compliance is needed to assess how well IRS is accomplishing its mission.

In light of the controversy surrounding IRS' last effort to measure voluntary reporting compliance by randomly auditing taxpayers, we initiated this review to determine the status of IRS' current efforts. In addition, we looked at how selected programs in other federal agencies use random sampling to measure compliance with the rules and regulations governing their programs. We undertook this review on our own initiative and, as agreed, are addressing it to you because of your interest in IRS' operations.

To determine the status of IRS' plans to measure voluntary reporting compliance, we reviewed available planning documents and discussed the plans with IRS officials. To obtain information on how others measure compliance, we identified six large programs in five other agencies that used random samples to gauge compliance. We then contacted and interviewed officials in the five federal agencies responsible for the six programs and reviewed available documentation. The six programs and associated agencies were the (1) Health Care Financing Administration's (HCFA) Medicare Program; (2) Department of Agriculture's Food Stamp Program; (3) Social Security Administration's (SSA) Old Age, Survivors, and Disability Insurance (OASDI) Program; (4) SSA's Supplemental Security Income (SSI) Program; (5) U.S. Customs Service's Compliance Measurement Program (CMP); and (6) Immigration and Naturalization Service's (INS) Inspections' Traveler Examinations (INTEX) pilot program. We did not attempt to assess the effectiveness of the agencies' random sampling programs. In appendix I, we discuss our objectives, scope, and methodology in detail.

---

We did our audit work from May 2000 through March 2001 in accordance with generally accepted government auditing standards.

---

## Results in Brief

In May 2000, IRS established the National Research Program (NRP) Office to develop an approach for measuring voluntary compliance. While IRS has not finalized plans for such a measure, the program staff has established study objectives and developed a set of guiding principles that are to be adhered to when conducting the study. The study objectives consist of gathering data to measure voluntary compliance and to design and enhance operating programs and services. The guiding principles for the study are to limit undue burden on taxpayers; seek design input from external stakeholders, including Congress and Department of the Treasury officials; and ensure that the data meet operational needs.

As of March 2001, IRS had a preliminary draft of a business plan for measuring voluntary compliance. Included in the plan are four alternatives related to measuring voluntary reporting compliance, including two that involve auditing random samples of returns; one that uses operational data; and one in which voluntary reporting compliance is not measured. Also, IRS was developing database software that is to interface with existing systems to collect and analyze data. IRS intends to conduct a comprehensive test of all the software because these systems are a key component of the program.

Each of the six programs we looked at measures compliance by gathering different types and amounts of information from a random sample of clients. Sample sizes range from about 1,400 to over 500,000 annually. In each program except Medicare, clients are randomly selected and interviewed in person. For the Medicare program, claims are sampled from providers of services, such as doctors and vendors providing medical equipment. Generally, Medicare officials conducting the reviews communicate with providers through correspondence or telephone. Officials from each agency told us that the reviews are necessary to determine the agency's progress in meeting its performance goals and to help managers identify areas where operational changes are needed.

In commenting on a draft of this report, the Commissioner of Internal Revenue agreed with our assessment of the NRP. He also noted that, subsequent to completion of our work, IRS proposed a somewhat modified approach that shifts even more of the burden to IRS and reduces the need to conduct traditional face-to-face audits. The Commissioner noted that IRS is beginning to seek external stakeholder help in refining

---

the approach to measuring voluntary compliance. The Commissioner's comments are discussed later in this report and are reprinted in app. IV.

---

## Background

One of IRS' strategic goals is to provide top-quality service to all taxpayers through fair and uniform application of the law. The Commissioner of Internal Revenue has stated that measuring voluntary compliance is essential for determining IRS' progress in meeting this goal.<sup>2</sup> However, IRS' last measure of voluntary reporting compliance was done on tax year 1988 returns, where an estimated 8 percent of taxes owed were underreported on individual tax returns. This measure was developed from data obtained under IRS' 1988 Taxpayer Compliance Measurement Program (TCMP), which audited a random sample of about 50,000 individual taxpayers.<sup>3</sup> IRS is concerned that these data are too old to be meaningful.

IRS last planned to measure voluntary reporting compliance for tax year 1994 returns under TCMP. The 1994 TCMP was an ambitious effort, with IRS planning to audit over 150,000 randomly selected tax returns of individuals, corporations, partnerships, and S-corporations. Over 92,000 of these audits were to be conducted on individuals (including sole proprietors and farmers) filing the Form 1040. Before beginning the audit process, however, IRS canceled TCMP because it was too costly.<sup>4</sup> Also contributing to the project's cancellation was criticism from Congress, the media, tax community, and taxpayers, because the sample was considered too large and the audits too burdensome on taxpayers.

We have previously reported that the results of TCMP studies provided more than just an indicator of compliance status. (See app. III for a list of the reports we have issued on TCMP studies.) The information generated from TCMP surveys was used to help IRS identify compliance trends, and it allowed IRS to suggest changes in tax laws and regulations to improve voluntary reporting compliance. TCMP results were also used to develop

---

<sup>2</sup>The National Commission on Restructuring the Internal Revenue Service also reported in June 1997 that it was important that IRS measure voluntary compliance.

<sup>3</sup>TCMP audits have been conducted periodically since 1963 on individuals, corporations, partnerships, S-corporations, and estates. Appendix II lists the TCMP surveys.

<sup>4</sup>IRS estimated that the 1994 TCMP would cost about \$559 million in direct costs over 3 fiscal years and about \$1.5 billion in opportunity costs, which are the revenues that would not be realized from conducting regular audits. While the primary purpose of TCMP audits was not to produce revenues, IRS estimated that the revenue yield from these cases would have been about \$685 million.

---

formulas that helped IRS more effectively identify noncompliant taxpayers for operational audits.<sup>5</sup> The formulas developed from TCMP data help IRS identify tax returns with a high probability that an audit would detect changes to tax liability. The lower the numbers of returns with no change to tax liability the better the formulas are as predictors of audit opportunity. According to IRS, since return selection formulas were first developed from TCMP data and used to select 1968 tax returns for audit, the number of audits that result in no change to tax liability have been reduced from about 46 percent in the year before the use of formulas to about 19 percent in fiscal year 1994. Starting in fiscal year 1995, however, the number of audits where there was no-change to tax liability after an audit began to increase, and it was 24 percent in fiscal year 1998, raising IRS' concerns about increasing the burden on compliant taxpayers.

TCMP data also enabled IRS to estimate the tax gap, which is the difference between the amount of tax owed and the amount of tax voluntarily paid. The last tax gap estimate was made using 1985 and 1988 TCMP data, in which IRS projected that the individual tax gap for 1992 would be over \$93.2 billion. TCMP data were also used by congressional tax writing committees, the Department of Commerce, Department of the Treasury's Office of Tax Analysis, and state governments. These uses included estimating revenue impacts of new legislation; making adjustments to the national income accounts, such as gross domestic product; and researching tax noncompliance. Some states have also used compliance measurement data to help them develop state compliance programs.

---

## IRS Has Not Finalized Its Plans on How It Will Measure Voluntary Reporting Compliance

In May 2000, IRS established a NRP Program Office to develop an approach for measuring voluntary compliance. While IRS has not finalized its plans for how to measure voluntary reporting compliance as of March 2001, some steps have been taken. The staff has established study objectives and developed a set of guiding principles that are to be adhered to when conducting the study. As of March 2001, IRS had a draft business plan that contained information on four different approaches for measuring voluntary reporting compliance, including two that use random audits. Also, in the event that IRS decides to use audits to measure compliance, the staff plans to use existing data collection software and has

---

<sup>5</sup>Operational audits are the audits of taxpayer's books and records done annually as part of IRS' enforcement effort. Returns are selected because IRS has some indication that the taxpayer is not compliant.

started developing database software to organize the data and make it more accessible to staff in the operating divisions. Given the importance of these systems to the program, IRS plans to conduct a comprehensive test of all the software.

## IRS Has Established Objectives and Guiding Principles for Developing a Measure of Voluntary Reporting Compliance

NRP is IRS' effort to measure its progress in meeting its strategic goal of providing top-quality service to all taxpayers through fair and uniform application of the law. The NRP Office has developed objectives and established several guiding principles to direct its efforts to plan and implement this program. Table 1 outlines these objectives and guiding principles.

**Table 1: Objectives and Guiding Principles for the NRP Program**

Objectives	Guiding principles
<ul style="list-style-type: none"> <li>• Measure taxpayer reporting compliance</li> <li>• Improve ability to detect taxpayer noncompliance and develop methods for allocating compliance resources</li> <li>• Support strategic measures and planning</li> <li>• Use results to improve tax system operations</li> </ul>	<ul style="list-style-type: none"> <li>• Minimize taxpayer burden as data are collected</li> <li>• Solicit external stakeholder ideas and support in design of the study</li> <li>• Engage IRS' operating divisions in the ownership of the study<sup>a</sup></li> </ul>

<sup>a</sup>IRS has four business operating divisions specializing in different types of taxpayers: Wage and Investment income, Small Business/Self-Employed, Large and Mid-Size Business, and Tax Exempt/Government Entities.

Source: IRS' draft business plan for measuring compliance.

The objectives for NRP cover a wide range of uses for compliance data. In addition to the objective of determining the voluntary reporting compliance rate, IRS expects the approved measurement approach to produce data that can be used to update return selection criteria so that it can better detect noncompliance, thus reducing the number of compliant taxpayers selected for operational audits. Similarly, NRP is intended to obtain data that can be used to improve IRS' operations by providing specific information about noncompliance, allowing IRS to make operational changes that could improve compliance. For example, TCMP data indicated that many taxpayers erroneously claimed dependents on their tax returns. IRS subsequently recommended that Congress require taxpayers to include the dependent's social security number on their returns. Taxpayers claimed 7 million fewer dependents on their 1987 tax returns, the first year dependent social security numbers were required.

---

The guiding principles represent the conditions under which the study will be conducted to meet the objectives. According to IRS, these principles were developed in response to certain real and perceived deficiencies with prior TCMP studies. For example, IRS recognized the need to limit the burden imposed on taxpayers by measuring their compliance. In 1994, taxpayer burden was one of the key criticisms of doing TCMP audits. Although the current plans for measuring voluntary reporting compliance are not complete, some proposed actions seem likely to reduce the burden imposed on taxpayers if IRS determines that random audits are necessary. These include (1) using a dedicated cadre of auditors for NRP audits, which IRS believes will reduce the time needed for an audit; (2) building a comprehensive preaudit case file for the taxpayers being audited, which will reduce the amount of information auditors must request from the taxpayer; (3) limiting the number of lines audited on some returns; and (4) minimizing the sample size, resulting in fewer taxpayers being inconvenienced by audits.

To avoid criticism like that received on its plan for the tax year 1994 TCMP, IRS planned to contact external stakeholders and solicit their ideas and support before making a decision on the final measurement approach. For example, at the time of our review, NRP staff had contacted officials at Treasury to begin discussing any concerns they might have with the study. The Commissioner indicated that he would be contacting interested congressional officials directly. Finally, NRP staff planned to contact members of the tax community, such as tax preparers and accountants. According to NRP staff, other stakeholders would be included as they are identified.

To ensure that the measurement approach meets the needs of the operating divisions, NRP officials have begun a dialogue with operating division staff. This process began when NRP staff briefed senior operating division managers on the program's status. Also, IRS has established an executive steering committee, chaired by the NRP Director, with representatives from the Commissioner's office, each of the operating divisions, and other potential stakeholders, to ensure high-level, oversight over the study.

---

## Components of IRS' Draft Business Plan for Measuring Voluntary Reporting Compliance

As of March 2001, the NRP Office had developed a preliminary draft business plan, providing information on how IRS intends to develop a measure of voluntary reporting compliance around the framework of the guiding principles. The plan also shows how IRS intends to determine the most appropriate method for meeting NRP objectives. The plan discusses the pros and cons of four approaches for measuring reporting compliance

---

and how they align with the program's objectives. These approaches are (1) do not measure voluntary reporting compliance, (2) use operational audit data (i.e., results of IRS' day-to-day regular audit program), (3) conduct annual random audits of a small sample of taxpayers, or (4) conduct audits of a larger sample once every 3 years. NRP staff has not yet completed their study of the costs and benefits of the various approaches. However, the alternative with a larger sample conducted once every three years is currently the recommended approach.

The draft business plan analysis of the approach to stick with the status quo and do nothing points out that this approach would not add to taxpayer burden, but also that it would not meet study objectives. The draft business plan analysis for the approach using operational data, while incomplete, suggests that it may be possible to measure voluntary reporting compliance in this manner. However, the plan notes that although statistical methods exist for using nonrepresentative data to project to an entire population, the methods may not be well suited to analyzing the operational audit data available to IRS. According to NRP officials, more detailed analysis is being done on this approach.

As part of its analysis of the two approaches using random samples, IRS plans to minimize the sample size to address the concerns raised about the proposed tax year 1994 TCMP sample being too large. IRS proposes to limit the sample size by reducing the number of groups it plans to analyze. The sample size may also be reduced because IRS does not plan to update the operational audit selection criteria for most nonbusiness taxpayers.<sup>6</sup> Finally, in previous compliance measurement studies, every line of a selected tax return was audited. In keeping with the guiding principle of reducing taxpayer burden, the draft business plan indicates that IRS does not intend to audit each line of every return selected for audit. According to NRP officials, criteria for when to audit a line on the tax return and how to analyze lines that were not audited have not been determined.

IRS also intends to include in its NRP business plan a detailed data analysis plan that will show how the information gathered in the study is to be used to meet the study objectives. In the past, we have emphasized the importance of having an analysis plan to show how the data are to be

---

<sup>6</sup>Generally, two factors determine the size of the NRP sample—number of returns needed to get a precise estimate of the compliance rate and number of returns needed to develop the audit selection formulas. In most cases, it requires more audits to develop the audit selection formulas than to precisely estimate the compliance rate.

---

used. Also, such a plan is important to ensure that institutional systems and staff are in place to quickly and efficiently analyze the data and get the results into the hands of managers who can make appropriate changes. NRP officials told us that they are working on these plans with operating division staff and that analysis plans will be included in the final business plan.

If either of the two random audit approaches or the use of operational audits is deemed the appropriate method for meeting NRP objectives, IRS plans to use existing reporting and data gathering systems to collect the data. NRP officials believe that using these existing systems—the Report Generating System and the Examination Operational Automation Database—will reduce the time needed to complete data entry. They also believe that using the existing software to collect NRP data will result in more accurate data being added to the database because audit information will be directly entered by auditors in an electronic format, with no need to transcribe it after the audit. IRS is also developing database prototype software that is to interface with the existing systems. IRS officials indicated that when the database program is completed, they plan to conduct an operational test of the combined systems to ensure that they are operating efficiently and will provide the services required for NRP.

---

## How Other Programs Measure Compliance

The six federal programs we studied, like IRS' TCMP studies, randomly sample their clients to measure compliance with the rules and regulations covering their programs and services. Agencies responsible for these six programs measure compliance to conform to the Government Performance and Results Act (GPRA) of 1993, which required that agencies establish measures to determine the results of agency activities.<sup>7</sup> In general, the agencies use random sampling to reduce the cost and burden of collecting this information. Each of the six programs we studied conducted compliance reviews differently, gathering different amounts and types of information with varying amounts of time required for the review. Table 2 summarizes the characteristics of the random samples used to measure compliance for each of the six programs.

---

<sup>7</sup>GPRA seeks to shift the focus of government decision-making and accountability away from a preoccupation with the activities that are undertaken—such as grants dispensed or inspections made—to a focus on the results of those activities, such as real gains in employability, safety, responsiveness, or program quality. The act requires federal agencies to set goals, measure performance, and report on their accomplishments.

**Table 2: Compliance Review Characteristics for Six Federal Programs**

Compliance review characteristic	Program					
	Medicare	Food stamps	OASDI	SSI	CMP	INTEX <sup>a</sup>
Sample size and units of measure	120,000 claims	55,000 clients	1,440 clients	7,000 clients	558,000 travelers	17,000 travelers
Size of sampled population (millions)	900	7.9	44.5	6.6	160.5	208.5
Sampling period for population	Annual	Annual	Annual	Annual	Annual	Annual
Meet with clients in-person	No <sup>b</sup>	Yes	Yes	Yes	Yes	Yes
Average time spent with clients (minutes) <sup>c</sup>	Not applicable <sup>d</sup>	40	45	60	15	17
Third-party contacts	No	Yes	Yes	Yes	Yes	Yes

<sup>a</sup>According to INS officials, this is a pilot program.

<sup>b</sup>Medicare compliance reviews are conducted on claims and only providers, not Medicare recipients, are contacted.

<sup>c</sup>Indicates time spent with client during the inspection or examination. Does not include pre- or post-interview time.

<sup>d</sup>Providers are generally contacted only through correspondence.

Source: Discussions with program officials and review of documents.

While none of the agencies' mission or responsibilities is exactly like IRS', there are similarities in how they and IRS treat compliance with the laws and regulations governing their programs. For example, the agencies and IRS try to ensure compliance through efforts that apply to all their clients. Similarly, each agency uses random samples to measure the effectiveness of these efforts and to determine how well they are accomplishing their mission objectives.<sup>8</sup> Another similarity, is that most of the agencies, when measuring compliance, contact clients in person. Finally, the agencies statistically project sample results to the entire client population.

Officials in each of the agencies we contacted indicated that the results of random sampling were used to measure progress against the agency's strategic goals. Most of the agencies also used the data to help make decisions about operational changes that could improve compliance. For example, HCFA used its compliance measurement data to track claims through the processing system and record their disposition so that it could identify areas of concern and develop corrective actions to improve the efficiency and accuracy of the claims processing system. Also, HCFA officials considered the analysis of data generated from compliance

<sup>8</sup>IRS' efforts to ensure compliance include taxpayer education and assistance, math error checks, information returns, and operational audits.

---

measurement to be a tool for assessing the validity of complaints received from health care providers as well as a method for evaluating the contractors providing payment services to the Medicare Program. INS officials told us that they may be able to use compliance measurement data to identify areas where additional training was needed for their inspection staff. Finally, Customs officials indicated that compliance reviews using random samples allow them to make estimates of how much contraband they fail to interdict on the primary inspections. According to these officials, without some secondary review using random sampling, they would only be able to provide Congress with data on the amount and number of illegal imports they were able to identify in their primary inspections. (A more detailed discussion of each agency's compliance review program can be found in app. IV.)

By using statistically valid random samples, agencies can obtain compliance information from a small fraction of the program's population. For example, as shown in table 2, 55,000 food stamp recipients are sampled out of 7.9 million total recipients. Random sampling allows state and federal officials to obtain data on overall compliance, but burden less than 1 percent of the program's population.

---

## Agency Comments

We obtained written comments on this report in a June 11, 2001, letter from the Commissioner of Internal Revenue (see app. IV). In his letter, the Commissioner noted that our report recognizes IRS' need to measure compliance and develop approaches to effectively identify noncompliance. He further noted that the NRP approach is continuing to evolve and has been somewhat modified to shift more of the burden away from taxpayers and onto IRS. In his letter, the Commissioner stated that the approach currently being considered maximizes the use of data available to IRS and further reduces the need to conduct traditional face-to-face audits.

The Commissioner emphasized that the success of NRP is dependent on the acceptance and support of the methodology by stakeholders. He noted that IRS is beginning to get external stakeholders involved in refining the NPR design and that the process will not in any way be considered final until this step is complete.

---

We are sending copies of this report to the Chairman and Ranking Minority Member, Subcommittee on Oversight, House Committee on Ways and Means; the Secretary of the Treasury; and the Commissioner of Internal Revenue. The report is also available at [www.gao.gov](http://www.gao.gov).

---

If you or your staff have any questions, please contact Ralph Block at (415) 904-2150 or me at (202) 512-9110. Key contributors to this assignment were Lou Roberts and Kathleen Seymour.

A handwritten signature in black ink that reads "James R. White". The signature is written in a cursive style with a large, prominent initial "J".

James R. White  
Director, Tax Issues

---

# Appendix I: Objectives, Scope, and Methodology

---

## Objectives

We initiated this assignment as part of our effort to monitor IRS' activities related to its reorganization, implementation of its new strategic goals, and development of measures to assess its progress in meeting these goals. Also, we wanted to explore how other agencies assessed compliance with their rules and regulations and measured progress against their strategic goals.

Our objective was to determine the status of IRS' plans to measure voluntary reporting compliance. In addition, we looked at how compliance is measured in six other federal programs.

---

## Scope and Methodology

To determine the status of IRS' efforts to measure voluntary compliance, we reviewed available planning documents, including a preliminary draft of the business plan. We discussed potential methodologies IRS might use to measure voluntary compliance with NRP officials. We also discussed potential sampling methodologies with various officials from IRS' Research Division as well as NRP officials. We reviewed preliminary documents related to the random sampling methodologies, including an assessment of IRS' documentation for the additional cases that would be required to update return selection formulas. We talked to officials responsible for developing and managing the data gathering and report writing systems IRS plans to use during its study. Additionally, we reviewed documentation for these systems.

We did not evaluate IRS' overall efforts to measure voluntary compliance because parts of the business plan had not yet been completed. We plan to continue to monitor and report on IRS' efforts to develop an acceptable approach for measuring voluntary compliance.

To describe how other agencies measure compliance, we reviewed information available at federal agency Web sites and identified five agencies and six programs where random sampling was used to determine compliance. We contacted officials from these agencies and reviewed information from the random sampling programs. Table 3 lists the agencies and the programs they administer.

---

**Table 3: Agencies Contacted and Programs Administered**

<b>Agency</b>	<b>Program administered</b>
Health Care Financing Administration	Medicare
U.S. Department of Agriculture	Food Stamps
Social Security Administration	Old Age, Survivors, and Disability Insurance Program
	Supplemental Security Income Program
U.S. Customs Service	Compliance Measurement Program
Immigration and Naturalization Service	Inspections' Traveler Examinations Pilot Program

Source: Agencies contacted.

At each agency, we obtained documentation about the random sampling efforts and discussed random sampling procedures with agency officials. We also discussed how the data from the random samples were used and how agency operations had been changed based on these data. Our work on the six programs was descriptive, and we did not attempt to assess the effectiveness of the random sampling programs.

# Appendix II: TCMP Surveys

The following table summarizes IRS' efforts to measure voluntary compliance using the TCMP surveys between 1963 and 1988.

**Table 4: Types of TCMP Surveys by Return Type and Tax Year**

Type of TCMP survey	Tax year	Sample size
Individual income tax	1963	92,000
	1965	50,000
	1969	53,000
	1971	26,000
	1973	55,000
	1976	50,000
	1979	55,000
	1982	50,000
	1985	50,000
	1988	54,000
Small corporations	1969	16,000
	1973	20,000
	1978	33,000
	1981	33,000
	1987	19,000
Estate returns	1971	4,600
Exempt organization returns	1974	11,400
	1979	20,000
	1988	3,000
Fiduciary returns	1975	8,900
Employee plan returns	1982	18,000
Partnership returns	1982	27,000
S corporation returns	1985	10,000
Delinquent returns—non farm business	1963	27,000
	1966	114,000
	1969	70,000
Delinquent returns—individual	1979	25,000
	1988	25,000
Surveys of delinquent accounts	1963	178,000
	1964	166,000
	1969	1,800,000
	1970	1,800,000
	1971	1,800,000
	1981	1,800,000
	1984	1,800,000

Source: *Tax Administration: IRS' Plans to Measure Tax Compliance Can Be Improved* (GAO/GGD 93-52, Apr. 5, 1993).

# Appendix III: Summary of GAO Assignments Related to Measuring Voluntary Compliance

GAO report	Summary
<i>Tax Administration: IRS' Plans to Measure Tax Compliance Can Be Improved</i> ( <a href="#">GAO/GGD-93-52</a> , Apr. 5, 1993)	Summarizes uses of TCMP data and outlines who uses the data. Identifies weaknesses of proposed changes and establishes criteria for evaluating proposed changes to measures of voluntary compliance.
<i>Tax Compliance: Status of the Tax Year 1994 Compliance Measurement Program</i> ( <a href="#">GAO/GGD-95-39</a> , Dec. 30, 1994)	Summarizes IRS' plans for the 1994 TCMP and discusses promising changes. Identifies several weaknesses in the plan that IRS needs to fix before implementing the project.
<i>Tax Compliance: 1994 Taxpayer Compliance Measurement Program</i> ( <a href="#">GAO/T-GGD-95-207</a> , July 18, 1995)	Testimony on 1994 TCMP before the Subcommittee on Oversight, Committee on Ways and Means. Discusses uses of TCMP data and status of planned 1994 TCMP effort. Discusses some of the criticisms of TCMP. Identifies GAO reports where TCMP data were used.
Letter to the Commissioner on TCMP Errors ( <a href="#">GAO/GGD-95-199R</a> , July 19, 1995)	Summarizes errors in audits for 1988 TCMP and suggests changes to codes to be used to categorize the cause of noncompliance for the planned 1994 TCMP project.
<i>Tax Administration: Information on IRS' Taxpayer Compliance Measurement Program</i> ( <a href="#">GAO/GGD-96-21</a> , Oct. 6, 1995)	Follow-up on issues raised in our December 1994 report concerning timeliness and the types of data IRS planned to gather for TCMP audits. Also, briefly discusses other sources of data on voluntary compliance and the relevance of TCMP data for alternative tax system proposals. Indicates how IRS responded to our recommendations.
<i>Tax Administration: Alternative Strategies to Obtain Compliance Data</i> ( <a href="#">GAO/GGD-96-89</a> , Apr. 26, 1996)	Summarizes the problems caused by cancellation of the 1994 TCMP project. This report also identifies sampling strategies that will reduce the sample size and still provide some data.

Source: Developed from prior GAO reports on TCMP.

---

# Appendix IV: How Compliance Is Measured for Six Programs

---

Similar to IRS' prior studies on measuring taxpayer compliance, some federal agencies randomly sample their clients to measure compliance with the rules and regulations covering their programs and services. In general, agencies use random samples to reduce the cost and burden of collecting information on how well the agency is performing its functions. We tried to focus on the following six areas, when the data were available: (1) sample population from which the sample is drawn; (2) sample size; (3) methodology used to select the sample; (4) what data were collected and how they were used; (5) length of time to conduct the case reviews; and (6) cost of the reviews.

The agencies we contacted have augmented their efforts to measure compliance since passage of the Government Performance and Results Act (GPRA), which required that agencies establish measures to determine the results of agency activities. GPRA seeks to shift the focus of government decision-making and accountability away from a preoccupation with the activities that are undertaken—such as grants dispensed or inspections made—to focus on the results of those activities, such as real gains in employability, safety, responsiveness, or program quality. GPRA also requires federal agencies to set goals, measure performance, and report on their accomplishments.

Measuring compliance is not the same thing as ensuring compliance. Agencies use various processes, including verification of submitted information, in an effort to ensure that all clients are in compliance with the rules and regulations governing the programs. For example, in the Food Stamp program, states verify income and household circumstances during the application process. In contrast, agencies measure compliance to determine the effectiveness of these efforts to ensure compliance. A key difference between efforts to assure compliance and efforts to measure compliance is that assurance generally applies to every client the agency serves while measurements are generally based on a random sample of a small portion of the clientele. The following sections describe the steps five agencies are taking to measure compliance with the laws and regulations governing the six programs we reviewed.

---

## HCFA, Medicare Program

The Health Care Financing Administration, part of the Department of Health and Human Services, administers the Medicare Program. Medicare annually pays more than \$200 billion to 1 million health care providers for services provided to nearly 40 million seniors and disabled Americans.

In August 2000, HCFA developed and began to implement a new measure that looks at the accuracy and documentation for payments to providers. This new compliance measure, called the Comprehensive Error Rate Testing (CERT) Program, is designed to assess how accurately providers bill the Medicare Program and improve on the accuracy and usefulness of existing studies.<sup>1</sup> HCFA has contracted the CERT Program functions to an independent third party that provides the sampled claims database, medical claim reviewers, and various management reports. The annual cost of this contract is \$4.1 million.

Over the course of each year, a random sample of 120,000 claims is chosen from the over 900 million Medicare claims received. Analysts randomly select 2,000 claims from each of 60 contractors who process Medicare claims—allowing HCFA to make contractor-specific compliance estimates. Medical records for the sampled claims are requested from the provider, and medical professionals independent of HCFA review the records for appropriateness of treatment and accuracy of billing. If necessary, the provider may be contacted and asked to explain items on the claim and related medical records. HCFA is phasing in the CERT project, starting with durable medical equipment claims. It plans to have a baseline error rate for the entire program by December 2002. Subsequently, the error rate will be updated annually.

According to HCFA officials, a primary use of the CERT data will be to measure how well they meet the goals established for GPRA. HCFA officials also noted that the data would be used to determine how well contractors were fulfilling their processing functions. Finally, HCFA officials told us that they believe that CERT data can be used to identify problems with the program, such as assertions by providers that claims they submit are often lost or misplaced by HCFA contractors.

---

## Department of Agriculture, Food Stamp Program

The Food Stamp Program is the largest of the domestic food and nutrition assistance programs. Administered by the U.S. Department of Agriculture's Food and Nutrition Service, food stamps worth \$16.9 billion were issued to an average of 19.8 million persons per month for fiscal year 1998.

---

<sup>1</sup>HCFA's Office of the Inspector General also conducts audits that provide a national payment error rate. Inspector General auditors select a statistically valid sample using a multistage, stratified sample design. This method estimates the extent of fee-for-service payments that do not comply with Medicare laws and regulations.

Each year, states identify a random sample of individuals receiving food stamps; the sample totals about 50,000 nationwide. States are responsible for conducting eligibility determinations for this sample of households, and certifying the correctness of these eligibility determinations. For example, the state of California had 865,632 food stamp recipients for which the state completed 1,103 reviews for fiscal year 1998. After the state reviews, the quality control group in the Food and Nutrition Service selects a subsample of cases reviewed by the states in order to assess the accuracy of these certification actions. These subsamples are also used to identify the state's error rate, which affects the amount of funding the state receives for administering the program. In 1998, the Food and Nutrition Service quality control sample ranged from 150 to about 400 cases per state, depending on the state's level of program participation. For example, 377 cases were sampled from California.

State reviews consist of face-to-face interviews with food stamp recipients to determine eligibility and whether they are receiving the proper amount of food stamps. Interviews are conducted at the client's home, or workplace. The client must provide documentation of their household circumstances, including information on each household member, their income, and the household's income. State reviewers also obtain information from external databases and other third parties. The interview generally lasts from 30 to 40 minutes. Data are collected on a 16-page data collection instrument. These data are aggregated, and a statewide error rate is calculated. The Food and Nutrition Service validates this process. Food and Nutrition Service officials estimate that the entire compliance and quality control program costs about \$72 million annually.

According to agency officials, these reviews are used to determine the quality of the states' efforts to screen applicants for food stamps and to determine how enhanced funding is awarded and liabilities are assessed. These reviews are also used to determine how effectively the program is meeting its goal of improving the accuracy of eligibility determinations and food stamp allotment amounts.

---

## SSA, Old Age, Survivors, and Disability Insurance Programs

The Social Security Administration (SSA) administers the Old Age, Survivors, and Disability Insurance (OASDI) Program, commonly known as Social Security. In fiscal year 1999, SSA made payments of over \$382 billion to more than 44 million claimants of retirement and disability insurance payments. With payments this large, even very small payment discrepancies can result in the loss of very large amounts of money. Given the potential magnitude of even small payment discrepancies and the need

to protect the Social Security Trust Fund, SSA has made improving payment accuracy a key strategic goal.

The payment accuracy rate is determined through an annual review of statistically valid samples in both the old age and disability programs. For the old age program, SSA randomly selects 80 cases per month (960 per year) for review of payment accuracy. In the disability program, 40 cases per month are selected (480 per year). According to SSA officials, these are statistically valid samples from which payment accuracy estimates can be made for the entire program population.

SSA reviews the case file and also contacts the claimant either in person or by telephone. Information is obtained from review of SSA files and from the claimant to assess the accuracy of payments. Claimant data are summarized on a 16-page questionnaire. Data obtained from the claimant include other names and social security numbers used in reporting earnings, marital history, disability information, and information on military service.

SSA officials told us that the samples and reviews are conducted primarily to measure how well the agency is meeting its objectives of improving payment accuracy. According to SSA officials, however, information from these samples is also used to identify problems with processes related to payment accuracy.

---

## **SSA, Supplemental Security Income Program**

The SSI Program is the nation's largest federally funded cash assistance program for the poor. The program is administered by SSA, and in 1999, paid about \$28 billion in benefits to 6.6 million recipients. SSI is a means-tested program designed to provide or supplement the income of aged, blind, or disabled individuals with limited income and resources.

SSA's Office of Quality Assurance and Performance Assessment annually assesses payment accuracy in the SSI Program using a random sample of program participants. SSA selects about 580 SSI recipients a month—almost 7,000 per year—to assess payment accuracy. The only requirement for selection in this review is that the recipient has received a SSI payment during the selection month.

SSA uses a 26-page case analysis worksheet to collect information on each of the selected SSI recipients. In each case, SSA conducts a desk review and also contacts the SSI recipient either face-to-face in their homes or over the telephone. During recipient contacts, SSA obtains demographic

information, documentation of marital status, work history, income information, and changes in living arrangements, including household composition and expenses. According to SSA officials, this worksheet requires about 1 hour to administer.

SSA uses the data obtained from these SSI reviews to calculate a payment accuracy estimate. This estimate is used to assess how effectively and efficiently SSA performs its day-to-day business processes and service delivery functions. These reviews provide the basis for reports to Congress and other monitoring authorities and measure payment accuracy for the SSI Program.

---

## Customs Service, Compliance Measurement Program

Customs is responsible for ensuring that all goods entering and exiting the United States do so in compliance with all U.S. laws and regulations. Customs inspects goods and persons arriving by land, sea, and air at over 300 ports of entry. During fiscal year 2000, Customs processed about 482 million passengers arriving at land and airport ports of entry.

Customs conducts several types of random audits as part of its management and quality control efforts. The largest of these random inspection programs is conducted at U.S. land border crossings, where during fiscal year 2000, Customs randomly selected over 326,000 privately owned vehicles entering the country for a detailed secondary inspection. In these inspections, Customs looks for illegal contraband, such as drugs, and other goods that were not declared.

Similar random inspections are conducted at the top 20 commercial air passenger ports of entry, where in fiscal year 2000, over 232,000 passengers were randomly selected for detailed secondary inspections. According to Customs officials, these secondary inspections range from 5 minutes to 20 minutes. Longer inspection times occur at land ports of entry, where the inspector must conduct a comprehensive inspection of vehicles. A shorter period is needed at airports, where inspectors generally only look at the contents of baggage. Inspectors complete a one-page data collection instrument, obtaining information such as country of origin, passport information, and other demographic data. Customs also uses a law enforcement network as a data resource for determining criminal activities.

Customs uses the information from these secondary inspections to provide an estimate of the number of violators and the effectiveness of its primary inspection process. These inspections also serve as a measure of

how well Customs is meeting the goals it established to comply with GPRA. Results of primary inspections only provide data on the number of interdictions made—and not the number that are missed. Customs officials indicated that using random sampling for secondary inspections better meets GPRA goals because the secondary inspections allow them to estimate how much contraband they failed to interdict on the primary inspections. Consequently, using random samples helps them provide Congress with better inspection data overall.

Customs officials indicated that the information gathered from these secondary inspections is used to develop training programs for inspectors and to indicate where changes in inspection programs are needed.

---

## INS, Inspections' Traveler Examinations Program

The mission of the Immigration and Naturalization Service is to ensure that individuals seeking entry to the United States are legally entitled to do so. Economic, demographic, and political pressures worldwide have resulted in continued growth in the number of international migrants seeking to come to the United States, both legally and illegally. INS conducted over 497 million primary inspections in fiscal year 1999 at more than 300 ports of entry. Each year since fiscal year 1999, as part of a pilot program, INS conducts about 17,000 secondary inspections on randomly selected travelers entering the country through 22 of the ports of entry selected for analysis. These inspections are rigorous, supplementary inspections of travelers who have already been cleared through primary inspection. INS is currently evaluating this pilot program.

INS officials indicated that, under the pilot program, inspectors personally interview the traveler and review his or her passport and any other documents needed to legally enter the country. INS inspectors complete a one-page data collection instrument and collect data such as citizenship status, date of birth, country of residence, gender, and factors that led to a determination of inadmissible status (if appropriate). INS also conducts physical inspections of personal items, such as contents of luggage, pockets, wallets, and vehicles. Like Customs, INS also uses a law enforcement network as a data resource for determining criminal activities.

According to INS officials, each of these inspections, on average, requires an additional 11 minutes for each traveler. INS officials said that primary inspections take less than one minute. According to INS officials, the primary purpose of this pilot program is to determine whether or not these inspections provide a statistically valid performance measure to identify

---

the extent to which noncitizens are being incorrectly admitted to the United States at these 22 sites. This information could also be used to train agents and modify procedures to improve operations. INS officials also noted that the pilot was used to measure the results of agency operations in accordance with GPRA. No data were available on the cost of these activities.

# Appendix V: Comments From the Internal Revenue Service



COMMISSIONER

DEPARTMENT OF THE TREASURY  
INTERNAL REVENUE SERVICE  
WASHINGTON, D.C. 20224

June 11, 2001

Mr. James R. White  
Director, Tax Issues  
United States General Accounting Office  
Washington, DC 20548

Dear Mr. White:

Thank you for the opportunity to review the draft report to the Chairman and Ranking Minority Member of the Senate Finance Committee and the Chairman and Ranking Minority Member of the House Committee on Ways and Means, entitled *Tax Administration: Status of IRS' Efforts to Develop Measures of Voluntary Compliance*.

As noted in your draft report, the IRS recognizes the need to measure compliance, identify compliance trends, and develop approaches to more effectively identify non-compliant taxpayers. However, the data to meet these needs must be obtained differently than it was in the past. What we have learned from our past experience is shaping our proposed methodology.

The IRS established the National Research Program (NRP) office last year to explore and develop innovative and efficient approaches to measure reporting, filing, and payment compliance that are less intrusive and less burdensome to taxpayers. NRP is not just a compliance measurement program; it is a critical part of the overall IRS organizational transformation. The Service's strategic goals and balanced measures and how we measure them are integrated into the NRP objectives and guiding principles.

Although the February 2001 draft NRP business plan you reviewed described a recommended approach that would be less intrusive and less burdensome to the taxpayer, the IRS is now proposing a somewhat modified approach that I believe is superior to the one recommended in the draft plan. The new NRP approach shifts more of the burden of data collection from taxpayers to the IRS. The approach maximizes the use of data available to the IRS and reduces the need to conduct traditional, face-to-face audits of taxpayers. In addition, the audits, when used, will not be line-by-line nor as intrusive as in prior studies. The new NRP approach is in direct response to the concerns of taxpayers and stakeholders and reflects comments you have shared with us during the review process.

Consistent with the program's guiding principles, the success of the NRP is dependent on the acceptance and support of the methodology by stakeholders. The IRS is

2

beginning to engage external stakeholders in the design of the program in order to help us refine our approach. The design will not in any way be considered final until this critical input is obtained. As the NRP approach evolves, our business plan will be updated accordingly and provided for your review.

I appreciate your feedback and welcome your continued review and support of our efforts to measure voluntary compliance.

Sincerely,



Charles O. Rossotti



---

## Ordering Information

The first copy of each GAO report is free. Additional copies of reports are \$2 each. A check or money order should be made out to the Superintendent of Documents. VISA and MasterCard credit cards are also accepted.

Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

***Orders by mail:***

U.S. General Accounting Office  
P.O. Box 37050  
Washington, DC 20013

***Orders by visiting:***

Room 1100  
700 4<sup>th</sup> St., NW (corner of 4<sup>th</sup> and G Sts. NW)  
Washington, DC 20013

***Orders by phone:***

(202) 512-6000  
fax: (202) 512-6061  
TDD (202) 512-2537

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

***Orders by Internet***

For information on how to access GAO reports on the Internet, send an e-mail message with "info" in the body to:

[Info@www.gao.gov](mailto:Info@www.gao.gov)

or visit GAO's World Wide Web home page at:

<http://www.gao.gov>

---

## To Report Fraud, Waste, and Abuse in Federal Programs

***Contact one:***

- Web site: <http://www.gao.gov/fraudnet/fraudnet.htm>
- E-mail: [fraudnet@gao.gov](mailto:fraudnet@gao.gov)
- 1-800-424-5454 (automated answering system)