

00257



*REPORT TO THE JOINT COMMITTEE  
ON INTERNAL REVENUE TAXATION  
CONGRESS OF THE UNITED STATES*

*BY THE COMPTROLLER GENERAL  
OF THE UNITED STATES*



**How The Internal Revenue Service  
Selects Individual Income  
Tax Returns For Audit**

Department of the Treasury

Procedures for selecting individual income tax returns for audit generally protect U.S. taxpayers against abuse.

Examiners can sometimes request returns for audit without having to explain why they need them.

Taxpayers who pay more tax than they should are less likely to have their returns selected for audit than those who pay less than they should.

Not enough is known about why taxpayers do or do not voluntarily comply with the tax laws.

IRS needs to take steps to improve its audit plans and to make sure that the plans are followed.

GGD-76-55

NOV. 5, 1976



COMPTROLLER GENERAL OF THE UNITED STATES  
WASHINGTON, D. C. 20548

B-137762

To the Chairman and Vice Chairman  
Joint Committee on Internal  
Revenue Taxation  
Congress of the United States

This report, one of a series of reports in response to your Committee's request, addresses the Internal Revenue Service's planning process and its procedures for selecting individual income tax returns for audit.

Copies of this report are being sent to the Director, Office of Management and Budget; the Secretary of the Treasury; and the Commissioner of Internal Revenue.

A handwritten signature in black ink, appearing to read "James B. Stacks".

Comptroller General  
of the United States

## C o n t e n t s

		<u>Page</u>
DIGEST		i
CHAPTER		
1	INTRODUCTION	1
	The individual income tax	1
	IRS as administrator	1
	Scope of review	4
2	WHY YOUR RETURN MIGHT BE SELECTED FOR SERVICE CENTER AUDIT	6
	Selection of returns for service center audit	7
	Conclusions	11
3	WHY YOUR RETURN MIGHT BE SELECTED FOR DISTRICT OFFICE AUDIT	12
	Computerized selection process	12
	Non-DIF-selected returns	22
	Conclusions	24
	Recommendation to the Commissioner of Internal Revenue	25
	IRS comments and our evaluation	26
4	EVALUATION OF DIF EFFECTIVENESS	28
	Comparison of DIF with pre-DIF	28
	Comparing DIF with other computerized selection techniques and manual selection	30
	DIF compared with random and perfect selection	31
	1973 DIF formulas compared to previous formulas	32
	Correlation analysis of DIF scores with no-change rate	32
	Effect of classifiers on DIF results	34
	Evaluation of equity in return selec- tion process	34
	Tailoring TCMP and DIF to local needs	36
	Conclusions	37
	Recommendations to the Commissioner of Internal Revenue	39
	IRS comments and our evaluation	39

CHAPTER		<u>Page</u>
5	HOW IRS DETERMINES THE NUMBER AND TYPES OF RETURNS TO BE AUDITED--LONG-RANGE STRATEGY	42
	Voluntary compliance	42
	How IRS determines its long-range audit strategy	46
	Opportunities to improve the planning process	51
	Conclusions	55
	Recommendations to the Commissioner of Internal Revenue	56
	IRS comments and our evaluation	56
	Need for IRS to provide more information in support of its appropriation requests	58
	IRS comments and our evaluation	59
	Recommendation to the Congress	60
6	HOW IRS DETERMINES THE NUMBER AND TYPES OF RETURNS TO BE AUDITED BY DISTRICT OFFICES--ANNUAL PLAN	61
	National office	61
	Regional office	65
	District office	67
	Conclusions	68
	Recommendation to the Commissioner of Internal Revenue	68
	IRS comments and our evaluation	69
	Implementing the annual plan	70
	Conclusions	73
	Recommendation to the Commissioner of Internal Revenue	73
	IRS comments	73
	The annual plan as a quota	74
	Recommendations to the Commissioner of Internal Revenue	76
	IRS comments and our evaluation	76
 <b>APPENDIX</b>		
I	Letter dated August 10, 1976, from the Commissioner of Internal Revenue	78
II	Unallowable items	110
III	Reasons individual income tax returns are selected for district office audit	121

APPENDIX

Page

IV.	Technical materials relating to development of discriminant function techniques for selecting individual income tax returns for examination by the Internal Revenue Service	123
V	GAO's approach to statistical analysis of relationship of DIF scores and no-change rate	130
VI	Plan compared to accomplishments, fiscal year 1974	138
VII	Principal officials responsible for administering activities discussed in this report	142

ABBREVIATIONS

DIF	discriminant function
GAO	General Accounting Office
IRS	Internal Revenue Service
TCMP	Taxpayer Compliance Measurement Program

**BLANK**

COMPTROLLER GENERAL'S  
REPORT TO THE JOINT COMMITTEE  
ON INTERNAL REVENUE TAXATION

HOW THE INTERNAL REVENUE  
SERVICE SELECTS INDIVIDUAL  
INCOME TAX RETURNS FOR AUDIT  
Department of the Treasury

D I G E S T

Although no major problems were evident, the Internal Revenue Service can improve its procedures for selecting individual income tax returns for audit. Also, congressional oversight of the Service's audit activities would be enhanced by having the agency present the rationale behind application of its audit effort among the various classes of taxpayers.

WHY YOUR RETURN MIGHT  
BE SELECTED FOR AUDIT

Returns can be audited by IRS service centers, where taxpayers file their returns, or by local district offices, where taxpayers have most of their direct contact with IRS. Most returns are selected for audit by a computer or by persons other than the examiners who will be auditing them.

Criteria for selecting returns for service center audit are usually so specific that personal judgment is but a minor factor in the process. (See p. 11.) Some returns selected for district office audit are chosen randomly in connection with special tests; others are chosen because they have a special feature, such as having been prepared by an unscrupulous preparer. But most are selected because the computer or an individual has determined that the returns, in general, have good audit potential. (See p. 24.)

About 70 percent of returns audited by district offices are selected by a two-stage system. Returns first are "scored" as to their audit potential by a computer using sophisticated mathematical formulas. The highest scored returns then are manually screened to determine if an audit is warranted and, in most cases,

what items of income and deductions should be examined. GAO and IRS tests have determined that this system is effective, but these tests have concentrated on the use of the computer. Little has been done as yet to evaluate the effectiveness of the manual screener. (See pp. 28 and 41.)

Taxpayers who pay more taxes than they should are less likely to have their returns selected for audit under this system than are taxpayers who did not pay enough--primarily because it is difficult for the manual screener to identify those who have made overpayments. The mathematical formulas used to score returns are also biased against the overpayer. (See p. 34.)

To overcome this deficiency the Commissioner of Internal Revenue should direct IRS to measure the effect of the manual screener on the computerized selection system and determine ways to make sure that a representative number of returns involving overpayments are audited. (See p. 39.)

IRS told GAO it had taken steps to eliminate the bias in the mathematical formulas and planned to take others, directed at the manual screener, to better insure that a representative number of returns involving overpayments would be audited.

If IRS' plans are put into action, the manual screener's role in selecting returns for audit will be restricted, but he will continue to be responsible for determining the audit's scope. IRS apparently recognizes the need to measure the effect of the manual screener on the selection process and says it will consider ways to do it. (See p. 41.)

Since examiners usually do not select returns to be audited, there is little chance for abuse in the selection process. But one aspect of this procedure requires attention. A return can be selected for audit directly by an examiner if he determines that he needs to audit

--a return filed by a taxpayer for years other than the one being audited or

--a return filed by another taxpayer that may have a bearing on the return being audited.

To obtain such a return the examiner merely completes a requisition and indicates, by code, a general reason for wanting it. For example, code 40 means "prior year return" and code 50 means "partner." But the examiner does not have to provide additional explanation as to why he needs the return. (See p. 25.)

IRS believes that these codes sufficiently explain why the returns are being requested and that any questions about an examiner's need for a return can be asked by the supervisor before he approves the request.

GAO disagrees. There is no assurance that the supervisor will ask any questions and the codes alone do not explain to supervisors and other levels of management

--why the examiner wants the return,

--what he found in auditing the primary return that aroused his interest in a secondary return, and

--the significance of questions that the examiner wants to pursue on the requested return.

Answers to these questions are important if IRS wants to be sure that examiners are requesting returns for valid reasons. (See p. 26.)

The Commissioner of Internal Revenue should require examiners, when requesting specific returns, to explain on their requisitions why they need the returns so that the requests can be adequately evaluated. (See p. 25.)

IRS says it is making a comprehensive review of all its codes to insure that they are properly defined. GAO believes that in so doing, IRS should consider GAO's concerns.

IRS' AUDIT PLANS NEED TO BE  
IMPROVED AND NEED TO BE FOLLOWED

In developing its long-range audit plan for the 5 years ending with fiscal year 1979, IRS compared the value of three long-range audit plans--one designed to improve voluntary compliance with the tax laws, one designed to maximize the tax yield, and one designed to strike a balance between these two. It selected the balanced plan which called for improving compliance through increased audit coverage in classes of income where compliance was low, and for assigning remaining auditing staff to the rest of the classes on the basis of yield. (See pp. 46 to 51.)

While the long-range planning process is basically sound, it could be improved.

--IRS has virtually ignored the contribution of service center audits to the rate of compliance. These audits do affect compliance because the centers contact taxpayers about problems on their returns. By not taking such audits into account, IRS overestimates the staffing needs of district offices to meet compliance goals.

--IRS has done insufficient research to identify factors affecting taxpayer compliance. Preliminary research has indicated that audits may not be the most critical factor, but IRS has not aggressively pursued this. (See p. 51.)

IRS prepares an annual nationwide plan for the number of returns to be audited. Portions of this plan then are allocated to regional offices and on down to district offices. Some district offices allocate the plan to groups of examiners within the district. (See p. 61.)

Allocation to the regions is based on their proportion of returns with the greatest probability of tax error--an indication of each region's relative compliance level. Thus, a region where compliance is apparently low would be allocated a larger portion of the planned audit work than one where compliance is apparently high. But IRS has to adjust its nationwide plan to account for imbalances between the number of audits that should be done and the audit staff available to do them. Thus, some taxpayers are audited or not audited merely because of where they live.

Regions and districts do not always follow these same procedures for allocating workload and adjusting for imbalances.

To justify a request to the Congress for more examiners in fiscal year 1974, IRS committed itself to additional audits and tax assessments. Then, to carry out these commitments, IRS deviated from its fiscal year 1974 plan and directed more audits of medium income taxpayers--historically the best compliers with the tax laws--and less audits of other classes of taxpayers known to be of lesser compliance. (See p. 70.) This deviation was inconsistent with IRS' long-range compliance goals.

Because the annual plan calls for a specific number of audits, it often has been equated with a quota system. While GAO saw no evidence of quotas for individual examiners, some examiners told GAO they felt pressured to complete audits and felt that this pressure prevented them from doing a quality job.

Some examiners apparently believe that they are being pressured to adhere to unreasonable time constraints and that IRS is concerned with quantity to the detriment of quality. GAO does not believe that an annual plan, in and of itself, is the problem. A realistic plan can provide for a specific number of audits without sacrificing quality. But because IRS prepares its plan based on what was accomplished in the past rather than what

can reasonably be accomplished in the future, GAO questions whether the plan is realistic. (See p. 74.)

The Commissioner of Internal Revenue should not commit IRS to a specific number of audits or amount of additional tax assessments to justify requests for more audit staff. The Commissioner should direct IRS to

- consider service center audits in developing its long-range plans,
- try harder to uncover factors affecting taxpayer compliance,
- insure that regions and districts develop audit work plans consistent with the national plan, and
- study the efficacy of time constraints imposed on examiners. (See pp. 56, 68, 73, and 76.)

IRS does not provide the Congress with complete information during the appropriation process to justify its budget requests for additional audit staff. It does not clarify what alternative long-range plans are available or why a particular plan was selected. IRS does not, for example, point out that its long-range plan calls for different rates of compliance at different levels of income.

Given (1) the need to assure equity in tax law administration and (2) IRS' previous deviation from its plan, the Congress should discuss with IRS its decisions regarding audit coverage. But the Congress cannot do that unless IRS provides it sufficient data. (See p. 58.)

Therefore, GAO recommends that the Congress request IRS to provide detailed information on its audit plans. This information should be provided as a part of IRS' annual appropriation request. (See p. 60.)

IRS does not agree that it should consider the impact of all service center audits on compliance in developing its long-range audit plan because

- most of these audits do not fall within IRS' definition of "audit,"
- it would be difficult to estimate the workload that these audits would generate in any given year, and
- IRS doubts that these audits, in total, have the same overall effect on compliance as do district office audits.

GAO believes that IRS' definition of audit is too restrictive for planning purposes, IRS is seeking unnecessary preciseness in its planning process by claiming that it would be difficult to estimate workload, and there is ample reason to believe the effect of service center audits on taxpayer compliance is substantial. (See p. 57.)

IRS plans to continue searching for economical ways to assess the factors affecting taxpayer compliance and says it will consider the impact of service center audits in any such assessment. (See pp. 56 and 58.)

IRS agrees that more uniformity is needed in developing workplans and that it should refrain from committing itself to a specific number of audits or amount of revenue in justifying its requests for additional audit staff. IRS does not agree that a controlled study is necessary to evaluate the reasonableness of the time constraints imposed on its examiners. It has an alternative approach, however, that should help alleviate the apparently unreasonable pressure being felt by some examiners to close cases but, in GAO's opinion, falls short of assuring reasonable time constraints. (See pp. 69, 73, and 76.)

**BLANK**

## CHAPTER 1

### INTRODUCTION

In a June 18, 1973, letter, the Joint Committee on Internal Revenue Taxation asked us to examine the policies and procedures established by the Internal Revenue Service (IRS) for auditing tax returns.

This is the first of two reports on individual income tax returns (forms 1040 and 1040A). In it, we review IRS' planning process and procedures for selecting returns for audit. The second report 1/ addresses the audit process and the adequacy of IRS' controls against unwarranted tax assessments.

### THE INDIVIDUAL INCOME TAX

Individual income taxation in the United States began in 1863, was declared unconstitutional in 1895, and resumed with ratification of the 16th amendment to the Constitution and enactment of the income tax law on October 3, 1913.

This tax is an important source of funds for Federal operations. Of about \$294 billion in Federal taxes collected in fiscal year 1975, individual income taxes accounted for about \$156 billion (53 percent).

Most income tax revenues are collected under the pay-as-you-go system whereby wage earners have money withheld from their paychecks. Self-employed persons make periodic tax payments directly to IRS.

### IRS AS ADMINISTRATOR

IRS strives, as administrator of the tax law, to encourage the highest possible degree of voluntary compliance--that is, the ability and willingness of taxpayers to accurately assess their taxes. IRS communicates the requirements of the law to the public, determines the extent and causes of noncompliance, and does all things necessary to enforce the law. Its enforcement activities include auditing returns, collecting delinquent taxes and penalties, and recommending prosecution of individuals who evade their tax responsibilities.

---

1/GGD-76-54 to be issued later.

Of all enforcement activities, IRS considers the audit of returns to be the greatest stimulus to voluntary compliance. According to IRS, of the \$1.69 billion appropriated to it for fiscal year 1976, about 36 percent, or \$604 million, related directly to the audit activity, including its prorated share of administrative support costs. Statistics on audits of individual income tax returns for fiscal year 1975 follow.

Number of returns filed in calendar year 1974	81,271,762
Number of returns audited	3,160,419
Recommended additional tax and penalties	\$1.4 billion

IRS audit and related activities are carried out by the national office in Washington, D.C.; 7 regional offices; 58 district offices; 10 service centers; the National Computer Center in Martinsburg, West Virginia; and the Data Center in Detroit, Michigan.

#### Service centers

The 10 service centers process tax returns and related documents using automatic and manual data processing systems and high-speed processing devices, maintain accountability records for taxes collected, and audit certain returns. The offices primarily concerned with processing and auditing individual income tax returns are:

- The receipt and control branch, which receives and sorts incoming returns, remittances, and taxpayer correspondence.
- The examination branch, which prepares returns for computer processing and extracts information from returns for audit and statistical programs.
- The input perfection branch, which resolves errors detected during computer processing.
- The data conversion branch, which transcribes, verifies, and corrects pertinent information on all tax returns and related documents.
- The computer branch, which processes tax information and documents for mailing to taxpayers and for internal use and which generates computer reports, statistical information, and other information used throughout IRS.

--The classification branch, which selects returns to be audited by the service center and maintains a system for (1) insuring that returns with the greatest tax potential are selected for audit and (2) reviewing the audit results.

--The correspondence audit branch, which examines and verifies the selected tax returns by corresponding with taxpayers. These audits, which involve less complex issues, are done by tax examiners.

### National Computer Center

The National Computer Center establishes, maintains, and updates the individual master file (a record of all individual income tax filers) through a large-scale computer system. All tax data and related information pertaining to individual income taxpayers are posted so that this file reflects a current record of each taxpayer's account.

The Computer Center receives information on filed tax returns from the 10 service centers. In addition to updating the master file, the Computer Center uses the information to determine each return's audit potential. Audit potential is determined through formulas programed into the computer. Using these formulas, the computer assigns weights to certain basic return characteristics and totals the weights to arrive at a score for each return. The higher the score, the greater the probability that an audit of that return will result in a significant tax change. This scoring process is referred to as the discriminant function (DIF) system.

### District offices

Under the direction of the national and regional offices, district offices administer districtwide programs for selecting and examining tax returns. A typical district audit division is composed of:

--An examination branch, which is staffed by revenue agents and/or tax auditors 1/ who are supervised by group managers. Revenue agents usually have a college education with a major in accounting. An accounting major is preferred because the agent is expected to resolve tax issues requiring a high degree of accounting

---

1/Where appropriate, revenue agents and tax auditors will be referred to collectively as examiners.

and auditing skills. Agents conduct their audits by interview usually at the taxpayer's home or at the taxpayer's or his representative's place of business. Generally, tax auditors have a college education or its equivalent but are not required to have any accounting or related business subjects. Before advancing to the journeyman level, however, they are required to have six units of accounting and are given IRS training in accounting and auditing techniques that enables them to examine most individual tax returns. They conduct their audits either by correspondence or by interview, usually at an IRS office.

- A returns program management staff, headed by a returns program manager, which develops and administers district programs for selecting returns for audit. Classifiers--examiners temporarily assigned to this staff--screen returns to determine their audit potential.
- A review staff, which reviews completed audits to assure that the examiner did a quality job and that the tax liability has been properly determined.
- A conference staff, which meets with taxpayers who disagree with examiners' findings and attempts to settle their disputes.
- A service branch, which maintains control over tax returns, types form letters and other correspondence to taxpayers, and performs other miscellaneous services.

In some districts, some of these functions and staffs may be consolidated.

#### Data Center

The Data Center generates statistical reports used by management to monitor audit activities and evaluate their effect on voluntary compliance.

#### SCOPE OF REVIEW

We examined IRS policies, procedures, and practices in selecting individual income tax returns for audit. We

- reviewed pertinent IRS records;

--interviewed classifiers, tax auditors, revenue agents, group managers, and other IRS personnel; and

--reviewed 1,516 randomly selected files on individual income tax audits closed in 1973.

We did our work at IRS' national office in Washington, D.C.; its Dallas, Philadelphia, and San Francisco regional offices; its Baltimore, Cheyenne, Los Angeles, and New Orleans district offices; and its Kansas City and Memphis service centers. The four district offices serve Maryland and the District of Columbia, Wyoming, the southern half of California, and Louisiana, respectively.

## CHAPTER 2

### WHY YOUR RETURN MIGHT BE

#### SELECTED FOR SERVICE CENTER AUDIT

Two questions that arise in most taxpayers' minds are: How does IRS select returns for audit? and will my return be selected?

If audit is broadly defined to mean any review of a tax return, then all individual income tax returns are audited. During initial processing, the service centers review all returns to determine, among other things, whether they contain deductions that appear unallowable under the law and to verify the taxpayer's computations. In fiscal year 1975, for example, the service centers computer-verified computations on about 81 million returns. As a result, 2 million taxpayers had their liabilities increased by \$270 million and 1.8 million taxpayers had their liabilities decreased by \$163 million. Also, all individual returns are evaluated by a computer to determine their audit potential.

If the definition of audit is limited to situations where IRS questions something on the return other than a mathematical error and gives the taxpayer a chance to support the questioned item, then only a small percentage of individual income tax returns are audited. In fiscal year 1975, for example, service centers audited 1,321,861 returns and district offices audited 1,838,558 returns or 1.6 and 2.3 percent, respectively, of the 81 million returns filed in calendar year 1974.

In this report, we use the limited definition of audit.

IRS uses an even more limited definition. According to IRS, an audit is made only when an examiner has to inspect a taxpayer's records--which definition excludes most of the service center audits included by our definition because, in most cases, the service centers do not have to inspect taxpayer records. Using its definition, IRS statistics show that it audited 1,941,042 individual returns in fiscal year 1975: 1,838,558 by district offices and 102,484 by service centers.

IRS points out that 1,941,042 examinations were conducted pursuant to 26 U.S.C. 7602 which authorizes IRS to examine books and witnesses. The other 1,219,377 audits are characterized by IRS as "limited contacts." IRS believes

these contacts should not be considered audits for management or planning purposes. For the reasons set forth in chapter 5, we disagree.

SELECTION OF RETURNS FOR  
SERVICE CENTER AUDIT

Most returns audited by service centers are selected for audit because they involve relatively simple and readily identifiable problems that can be resolved easily by mail. Such audits are made under several different programs.

<u>Program</u>	<u>Number of returns</u>	
	<u>Audited in fiscal year 1975</u>	<u>Percent of total (note a)</u>
Unallowable items	952,120	72
Head of household	209,405	16
DIF correspondence	67,259	5
Information returns	34,838	3
Multiple filers	9,765	1
Federal-State cooperative	9,076	1
All other	39,398	3
Total	<u>1,321,861</u>	<u>101</u>

a/Total exceeds 100 percent due to rounding.

Unallowable items program

IRS questions items on individual income tax returns, identified manually and by computer, which appear to be unallowable by law. Of about 81 million individual returns filed in calendar year 1974, the service centers questioned about 952,000 because of unallowable items--about 1 return for every 85 filed.

During 1974 the unallowable items program covered 51 items. 1/ For example, IRS instructions provided that a casualty or theft loss will be identified as unallowable if information on or attached to the return indicates that the taxpayer

--is claiming the full loss without deducting the first \$100,

1/For 1976 the number of unallowable items increased to 64.  
(See app. I.)

- is claiming more than one casualty loss without deducting \$100 for each loss,
- has added \$100 to his loss rather than deducting it, or
- is claiming losses which in themselves are unallowable, such as termite losses and lost but not stolen items.

In addition to the 51 items, the examination branch can identify other items that appear to be unallowable by law. For example, if an examiner notices that a form W-2 is for a fiscal year rather than a calendar year, he codes the return unallowable and attaches a note identifying the problem for audit personnel.

#### Head of household program

IRS identifies returns on which the taxpayers appear to have erroneously claimed the unmarried head of household tax rate, such as returns in which the taxpayer claimed only one exemption. These tax returns are identified by computer. IRS determines if the taxpayers are entitled to the unmarried head of household rate by sending them a short questionnaire requiring yes or no answers.

#### DIF correspondence program

Classifiers request the highest DIF-scored low- and medium-income nonbusiness returns, review them, select those to be audited by the service center or the district offices, and accept the remaining returns as filed. Returns are designated for audit by the service center if they involve problems that can be resolved easily by mail (such as interest and contributions) and if information on the return indicates that the taxpayer can communicate effectively in writing.

#### Information returns program

IRS enters information on certain types of income, such as wages, dividends, and interest, from a taxpayer's return into a computer for comparison with information filed by employers, banks, and dividend-paying establishments and with Social Security Administration wage information. From the comparison, an inventory of potential underreporters is generated, from which the audit division selects specific returns on the basis of tax potential. For each return selected, a computer-printed transcript is prepared, comparing

the amount and type of income reported on information documents with that reported on the taxpayer's return. The transcripts are then screened by service center audit personnel to further evaluate tax potential, and the returns of apparent underreporters are selected for classification. In classification, the tax returns are compared to the transcripts to identify deviations. If a determination can be made that the taxpayer reported all income but in the wrong place on the return or that the amount of the cumulative discrepancy is minimal, the return is accepted as filed. If the cumulative discrepancy is large, however, the return is sent to the correspondence audit branch for examination.

#### Multiple filers program

IRS identifies instances where more than one return has been filed for a single year under the same social security number. The National Computer Center checks returns to identify multiple filings and returns so identified are extracted for audit. Classifiers review the returns to determine whether they should be audited by the service center or by the district office. The basic criterion for this decision is the complexity of the case. For example, two joint returns filed by the same taxpayer will be selected for examination by the service center while two nonjoint returns involving duplicated dependency exemptions or deductions will be forwarded to the district office.

#### Federal-State cooperative audit program

Copies of examination reports from State tax agencies are referred to the service center for association with Federal returns. These reports and the associated Federal returns are reviewed by classifiers to identify the returns to be examined. If the State audit adequately covered all major issues, the return is selected for service center audit; if not, the return is referred to the appropriate district office for audit.

#### Claims

This program involves the verification of refund claims and amended returns filed by taxpayers, with issues that can be effectively handled by correspondence. Other claims are audited by the district office. Kansas City service center guidelines for this program, for example, provide, in part, that an evaluation be made of all documents in the file. If enough information is available to reasonably accept the

claim or if the claim is not worthy of examination, it is to be accepted. Also, if the item on the claim would not have been questioned on the original return, it is not to be considered questionable on the claim.

#### Social security referral

Social security forms OAR-7000 (Notice of Determination of Self-Employment Income) are referred to service centers when the Social Security Administration has made a determination of self-employment income. The referrals involve adjustments to tax returns for self-employment and, possibly, income tax. If the potential tax change is less than a specified minimum amount, the returns are accepted as filed, unless there are other audit issues.

#### Interest paid on redemption of H bonds

Federal Reserve Banks send copies of information reports on H bond payees to the service center for the district in which the bank is located. The taxpayer service division assembles these reports with the payees' tax returns and refers them for classification by audit personnel.

#### Highway use tax

Service centers received information, through a private organization, on State motor vehicle registrations. The information was matched with highway use tax returns to determine deficiencies. Returns with apparent deficiencies were sent to classifiers who screened out cases involving a large number of vehicles. Those cases were forwarded to the appropriate district office; cases involving a small number of vehicles were retained for service center examination. This service center program was discontinued in November 1975.

#### Runaway parents

State welfare agencies periodically requested the last known address of a parent who had deserted and no longer supported his or her family. In addition to supplying the requested addresses, IRS used the names and social security numbers provided by the welfare agencies to identify returns for audit by checking whether the runaway parent had claimed a spouse and/or children as exemptions. Only nonjoint returns were selected for examination by the service centers;

joint returns were referred to the appropriate district office. This service center program was discontinued as of November 1975.

- - - -

Not every return meeting the criteria for audit under one of the service center audit programs is selected for audit by the service centers. If, for example, a return meeting the audit criteria for the unallowable items program is found to contain other audit problems, it will be referred to the appropriate district office. A return will not be selected for audit if the classifier determines that the additional tax would be minor.

#### CONCLUSIONS

Returns selected for service center audit involve relatively simple and readily identifiable problems that can be resolved by mail. Many are selected because they have a special feature, such as an unallowable item or a questionable use of the head of household tax rate, that, in effect, flags them for audit.

We did not review each service center audit program in sufficient depth to enable us to comment on the adequacy of the selection procedures and practices. We can say, however, that the procedures adequately protect against abuse in the selection process. Most of the returns and audit issues are identified either by the computer or by persons totally separate from the audit function. Other returns and issues, such as those in the Federal-State cooperative audit program, are actually identified by agencies outside IRS. Also, with two exceptions, the criteria for selecting returns for audit by the service center are so specific that judgment plays only a minor role in the process. The two exceptions are the DIF correspondence and the claims programs that require classifiers to decide which returns and which issues on those returns should be audited. Even here, however, the decision is being made by someone other than the person who will be responsible for auditing the return.

## CHAPTER 3

### WHY YOUR RETURN MIGHT BE SELECTED

#### FOR DISTRICT OFFICE AUDIT

Most returns audited by district offices involve issues that are not as readily identifiable or as easily resolved as those audited by service centers. Some returns are selected because of some special feature, such as a preparer whom the Internal Revenue Service has reason to believe is unscrupulous. Most, however, are selected because IRS has determined, through the discriminant function system, that the return has good audit potential.

The various reasons why a return might be selected for district office audit are listed in appendix III and are segregated into the following six major categories, as illustrated for fiscal year 1975.

<u>Major selection categories</u>	<u>Number of returns audited</u>	<u>Percent of total returns audited</u>	<u>Percent of total additional tax and penalties recommended for assessment</u>
Computer selection (DIF) (note a)	1,266,103	68.8	37.2
Audits initiated by IRS and others	178,025	9.7	16.1
Claims and other requests for refunds	64,843	3.5	.9
Related pickups	91,298	5.0	16.5
Multiyear audits	88,145	4.8	11.3
Miscellaneous	150,144	8.2	18.0
Total	<u>1,838,558</u>		

a/DIF accounted for about 98 percent of the returns audited under this category.

#### COMPUTERIZED SELECTION PROCESS

Before 1962 individual tax returns were selected for audit by a manual screening process. This process had several shortcomings, including:

--A poor use of highly skilled audit personnel. Hundreds of experienced examiners were assigned annually to screen 15 to 20 million returns.

--A lack of uniform criteria in the selection process because of the large role played by professional judgment.

--An inability to consider all returns for audit because of the number of returns filed each year. For example, before 1962, 15 to 20 million returns were screened each year, about one-fourth of the number filed.

Thus, some returns with lesser tax changes were audited, and some returns with potentially greater tax changes never entered the screening process.

In 1962 IRS began screening tax returns by computer. This method made use of 38 to 50 different classification criteria, depending on the type of return. Most of the criteria were item ratios, such as the ratio of contributions to reported income. At first, any return meeting one of the criteria was identified for audit, and the number of returns identified greatly exceeded IRS' audit capacity. Experienced audit personnel were called on to manually screen the computer-identified returns and weed out those with less audit potential to arrive at a manageable workload.

The criteria identification approach was modified in 1966. Returns were ranked by the number of individual criteria each met, and those meeting the greatest number of criteria were identified for examination first. This approach was instituted on the theory that the greater the number of criteria a return met, the greater the audit potential. Audit results improved, and manpower requirements for manual screening were further reduced.

Although this system was an improvement over the manual process, it was still inadequate. Weighting factors could not be applied to the criteria used in the selection process; large numbers of examiners were still needed to perform secondary screening; data processing costs were relatively high; and only marginal success in identifying returns with high audit potential was achieved.

#### Development of the DIF system

To correct the problems experienced with these early computerized selection systems, IRS undertook a project to

develop an effective and economical method of selecting individual tax returns for audit.

The objective was to screen the 70 to 80 million individual returns filed each year and then rank them as to audit potential. IRS hoped to (1) reduce the amount of computer time and manpower required to screen the returns, (2) direct auditors to the more productive returns, and (3) reduce the likelihood of auditing returns that were correctly prepared.

This project involved using 1961 Audit Research Program data to test the feasibility of developing effective mathematical formulas to select returns for audit. The Audit Research Program, a forerunner of IRS' Taxpayer Compliance Measurement Program (TCMP), involved auditing a random sample of low income business returns. After this test proved successful, formulas were developed for all classes of individual returns using TCMP data. TCMP is a program for measuring and evaluating taxpayer compliance characteristics through specialized audits of randomly selected tax returns. The individual return phase of this program was begun in 1964 with a random sample of 1963 returns. Later TCMP cycles involved random audits of 1965, 1969, 1971, and 1973 returns.

Data from TCMP audits provided the basis for grouping returns by the amount of tax change after audit. Items on the tax returns (referred to as variables), such as adjusted gross income, filing status, other dependents, rents and royalties, depreciation, repairs, and other expenses, were mathematically related to one of two tax change groupings-- returns with little or no audit potential and returns with a high probability of tax increase. The variables which helped best to separate the returns into these two groups were weighted and the weighting factors become the basis for the returns selection process.

The statistical process which determines the appropriate weighting factors is called discriminant function analysis. During the initial stages of development, IRS attempted to assign the weighting factors using certain modifications of classical discriminant function analysis. (See app. IV for a discussion of the mathematical process involved.) IRS also contracted with a private firm to develop weights using a similar, but proprietary, technique. According to IRS, when the two weighting systems were evaluated they were found to be comparable, but IRS chose the contractor's method because at that time IRS had limited data processing support and skilled resources to develop all the formulas on a timely

basis. This method became what is now called the DIF system. We did not examine the contractor's methodology because of its proprietary nature, although discussions with IRS indicate that it does not differ significantly from the classical approach.

The DIF score of a return is the sum of the weighted variables, and the highest scored return has the greatest audit potential. The DIF system for individual returns became fully operational in 1970.

The process of assigning weights to certain return characteristics and scoring the return is accomplished through the use of mathematical formulas. At the time of our review, IRS had a formula for each of the following audit classes.

<u>DIF formula</u>	<u>Audit class (note a)</u>
1	Low nonbusiness--Standard (adjusted gross income less than \$10,000) excluding returns with interest and dividends of more than \$200, other income, or adjustments to income.
2	Low nonbusiness--Itemized (adjusted gross income less than \$10,000) including returns with standard deductions and with interest and dividends of more than \$200, other income, or adjustments to income.
3	Low business--Schedule C or F (adjusted gross income less than \$10,000).
4	Low business--Schedule F for Office of International Operations use only (adjusted gross income less than \$10,000).
5	Medium nonbusiness--(adjusted gross income of \$10,000 but less than \$50,000).
6	Medium business--Schedule C (adjusted gross income of \$10,000 but less than \$30,000).
7	Medium business--Schedule F (adjusted gross income of \$10,000 but less than \$30,000).
8	High nonbusiness--(adjusted gross income of \$50,000 or more).
9	High business--Schedule C (adjusted gross income of \$30,000 or more).
10	High business--Schedule F (adjusted gross income of \$30,000 or more).

a/To measure voluntary compliance and allocate its audit resources, IRS groups taxpayers into audit classes on the basis of adjusted gross income. The business classes are for taxpayers who attach a Schedule C (Profit and Loss from Business or Profession) or Schedule F (Farm Income and Expense) to their returns. These should not be confused with corporate returns. The nonbusiness classes are for all other individual taxpayers.

New DIF formulas used to score the 1973 individual tax returns differed from those previously used, in that (1) some of the new formulas were developed by IRS statisticians, (2) information obtained from the 1969 TCMP was used for the new formulas whereas the previous formulas were based on 1965 TCMP data, and (3) the new formulas allowed IRS to use items from prior years' returns, such as number of dependents, when scoring the current year's return.

IRS feels that, since DIF has been used to score tax returns, it has been possible to demonstrate that the higher the DIF score, the more frequently an audit results in a tax change. It should be noted that DIF is not a substitute for but, rather, an aid to manual screening. DIF provides a method of differentiating between returns with a high likelihood of tax change and those with a lesser likelihood. Returns identified as having a high likelihood of change are still screened by classifiers who determine--by reviewing the entire return, including supporting schedules, and using judgment based on experience--whether the returns warrant audit. Thus, through DIF, IRS tries to increase the frequency of auditing returns that need to be audited.

#### Returns selection methodology study

IRS is seeking ways to improve the DIF system and exploring other approaches to supplement or replace DIF.

The returns selection methodology study included (1) investigating and suggesting ways to improve older DIF formulas, (2) investigating ways to better compare and evaluate those formulas developed under contract with those developed by IRS, (3) examining the feasibility of a two-stage formula that combines discrimination with predictive capability, (4) developing formulas for two groups of districts so that district differences could be examined, (5) developing a model which would incorporate audit-corrected information from TCMP in DIF formulas, (6) evaluating potential predictive techniques--such as stepwise multiple regression--that would allow IRS to predict the size of the tax change from tax return characteristics, and (7) investigating the feasibility of applying other mathematical and statistical techniques in the returns selection process to overcome the effects of interrelationships between return characteristics and enhance the potential capability of regression techniques.

### Selecting returns by DIF

Returns selected for audit under the DIF system accounted for about 70 percent of all returns audited in fiscal year 1975. A detailed description of how these returns are selected for audit follows.

#### Service center operations

A service center, after receiving a return, processes it through several stages. Such information as the taxpayer's name and address, gross income, withholding credits, refund due, and specific information used in computing the DIF score, is entered into the computer.

From the data entered, the computer produces two magnetic tapes at the end of each day--a "good" tape and an "error" tape. On the error tape are items that are not correct; for example, the taxpayer may have added wrong or used the wrong tax table.

An error register, which is printed from the error tape, and the related returns are then sent to tax examiners who review each item on the register and track down the cause of the discrepancy. In most instances, the process involves referring back to the tax return. After correction, the data is put on a good tape which is then transferred to the National Computer Center.

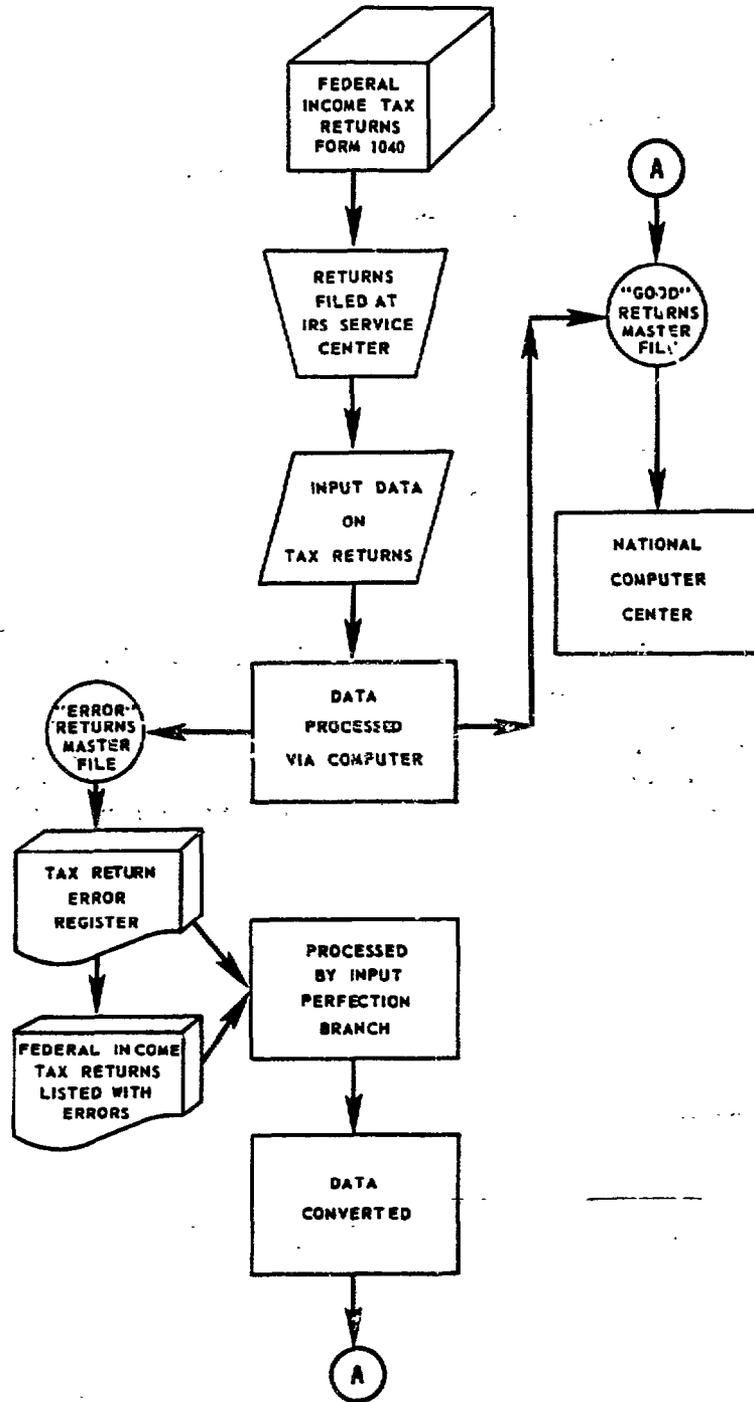
These service center processes are depicted in figure 1.

#### National Computer Center operations

Information received from service center tapes is used by the National Computer Center to score the returns. Each return is scored by the applicable DIF formula.

For each audit class, a minimum DIF score is established. Returns with scores at or above the minimum are listed on a DIF inventory file; returns with scores lower than the minimum are excluded.

After processing at the Computer Center, tapes from the DIF inventory files are shipped to the service centers. From the tapes the centers produce weekly inventory reports, which show the scores of DIF returns by audit class and post of



duty. 1/ These reports provide the district with the capability of ordering the highest scored returns.

#### District office operations

The district returns program manager periodically orders DIF returns from the service center. The returns are usually ordered by audit class and district and, where workload shortages exist, by post of duty. Returns with the highest DIF scores are sent to the district where classifiers (examiners temporarily detailed to the returns program management staff) manually screen them to segregate those in need of audit from those not in need of audit.

In other words, the computer, using formulas developed from TCMP data, evaluates the audit potential of every individual tax return filed and, by assigning scores, separates those with a high likelihood of tax change from those with a lesser likelihood. Then the classifiers, using judgment based on experience, evaluate the high-scored returns and determine which warrant audit and which do not. A return, for example, may have received a high score because of certain unusually large deductions. The classifier, however, upon reviewing the return, may see attached to the return what the computer could not see--detailed schedules in support of the deductions--and, in the absence of other questionable items, will determine that the return does not warrant audit.

Classifiers' decisions to select or reject returns for audit are subject to review by the returns program manager to insure that the classifiers exercised good judgment and arrived at their decisions in accordance with established procedures.

Most audits of DIF-selected returns are not comprehensive but are limited to certain questionable items on the return. The scope of audits of nonbusiness returns performed by tax auditors is determined by the classifiers. The scope of audits of business returns by tax auditors and all audits by revenue agents is determined by the examiners after a classifier has determined the return has audit potential.

---

1/Each district is divided into geographical segments served by local offices called posts of duty.

When screening nonbusiness returns, classifiers evaluate whether the income and deductions reported on the return would leave the taxpayer with enough money to cover necessary non-tax-deductible expenditures, such as for food and clothing. They also look for inconsistencies such as those between the type of employee business expenses claimed and the taxpayer's occupation, the expenses reported for real estate tax and home mortgage interest and the area where the taxpayer resides, and the asset value and asset life claimed for depreciation on rental income property. They also look for such things as misclassification of long- and short-term capital gains and losses; suspiciously rounded figures; and areas where past experience has shown taxpayers are more apt to make errors, such as bad debts.

On business returns, examiners look for the same items and inconsistencies as on nonbusiness returns, plus such matters as adequacy of the amount of gross receipts or cost of goods sold for the type of business conducted; reasonableness of business asset values and asset life claimed for depreciation; and unusually large amounts in accounts that might improperly include such personal expenses as legal fees, insurance, and travel and entertainment.

After a classifier selects a return for audit, he determines whether the examination should be made by a tax auditor or a revenue agent, depending on the complexity of the issues involved and the degree of accounting and auditing skills required to properly perform the audit. Less complex returns are audited by tax auditors; more complex returns are audited by the more highly qualified revenue agents.

Classifiers also prepare a checksheet for each return selected for examination by a tax auditor. These checksheets (1) indicate the particular items to be considered in the examination, (2) indicate whether the examination should be handled by correspondence or interview, (3) assist group managers in screening returns before assignment, and (4) assist in evaluating the classification program--the items classified are compared to the items adjusted after audit to evaluate the effectiveness of the classifiers in pinpointing questionable items. Checksheets can, at the discretion of the district office, be prepared for returns selected for examination by revenue agents.

#### Post of duty operations

Returns selected for audit by classifiers are sent, at regular intervals, to local offices called posts of duty.

Before a return is sent to the post of duty, a centralized group, such as the district office service branch, may notify the taxpayer that his return has been selected for audit and tell him to either bring or mail certain records to IRS. The return is then stored centrally and is not shipped to the post of duty until the taxpayer has mailed in his records or until immediately before the scheduled interview.

The group manager at the post of duty assigns returns to examiners on the basis of such factors as priority; geographical coverage; and examiner's experience, specialization, and/or workload. When the taxpayer has not already been contacted, the manager may screen the returns to (1) eliminate those he believes do not warrant examination and (2) insure that there is not some item on the return that should be considered during the audit but is not noted on the classification checksheet. If he believes a return warrants examination, he assigns it to an examiner who may again screen the return for audit potential if the taxpayer has not already been contacted. If the taxpayer has already been contacted, the return is not screened again at the post of duty because any changes as a result of that screening could cause taxpayer relations problems.

If, upon screening, the examiner determines that a return has audit potential, it becomes part of his inventory. If the examiner feels it does not have audit potential, he presents his reasoning to the group manager for final decision. If the manager agrees, the return is sent back to the service center; if he does not agree, the return is audited.

#### Pertinent statistics

Under the DIF system (computer scoring plus manual screening), all returns are evaluated for audit potential but relatively few are actually audited. For example, according to IRS:

--74.4 million individual income tax returns were processed and scored in 1972.

--Of these, 2.2 million were ordered by the district offices for classification.

--As of December 31, 1973, 0.7 million of these returns had been audited, 1.4 million had not been audited either because the classifier or group manager determined that they did not warrant audit or because they

could not be audited due to staffing limitations, and 0.1 million were either still in the audit stream or were unaccounted for.

#### NON-DIF-SELECTED RETURNS

In fiscal year 1975, about 30 percent of the returns audited by district offices were selected for audit for reasons other than their DIF scores. The reasons fall into five major categories.

##### Audits initiated by IRS and others

The return preparers program, which is directed against unscrupulous tax return preparers, accounted for 53 percent of all returns audited in this category and 5 percent of all returns audited by district offices. One IRS method for identifying questionable preparers is to have examiners prepare a report for the Intelligence Division whenever a return is adjusted for obviously unsupported items. As the Intelligence Division accumulates these reports, repeating preparers are noted and investigations are initiated to determine if they are guilty of fraudulent practices. Put simply, the returns selected for audit under this program are selected because IRS has reason to believe that the preparer of the return is unscrupulous.

Other returns included in this category are selected on the basis of information reports. Whenever an employee in the Audit Division receives information, from an audit or a third party, that a return filed or to be filed by a taxpayer will result in an additional or delinquent tax liability, he is to prepare an information report. This report, containing the source and nature of the information received, is submitted to the employee's immediate supervisor for approval. The supervisor is to approve the report only if it appears that the audit effort required would generate a material amount of additional or delinquent tax. The report is then associated with the affected return and both are reviewed by a classifier to determine whether the return warrants audit.

##### Claims and other requests for refund

When IRS receives a claim or other request for a refund or an adjustment in taxes, the original return may be manually screened to determine if the effort needed to substantiate the claim is warranted. In many cases, the audit results in disallowance of the claim or in assessment of additional tax and penalties.

### Related pickups

During an audit the examiner may find it necessary to review additional returns affecting the income and deductions of a taxpayer to ascertain whether the taxpayer correctly determined his liability. Included in this category are returns filed by partners, family members, and employers and employees of the taxpayer.

To obtain a related return, the examiner prepares a requisition which must be approved by his immediate supervisor. On the requisition he indicates, by a code number, why he is requesting the return. For example, one code indicates that the return is being requested because it was filed by the taxpayer's partner. The examiner is not required to provide any more information to justify his request. If the requisition is approved, the service center will forward the return directly to the examiner.

### Multiyear audits

During an audit the examiner may find it necessary to audit returns filed by the same taxpayer in earlier or later years to determine, for example, whether loss carrybacks or carryforwards are proper and whether adjustments to the return being audited might apply to other years' returns. To obtain the desired returns, the examiner follows the same procedures as previously described for related pickups.

### Miscellaneous

This category is the catchall for returns not specifically falling into the other categories. Of the returns covered by this category, the most important are those that are selected for audit under TCMP and those that are selected because of suspected fraud.

- - - - -

With some exceptions, non-DIF-selected returns go through a manual screening process, similar to the process previously discussed for DIF-selected returns, to determine whether they should be audited. The primary exceptions are (1) returns selected for audit under TCMP which are randomly selected and automatically audited and (2) returns, audited as related pickups or as part of a multiyear audit, which are sent directly to the examiner who requested them.

## CONCLUSIONS

Some returns are randomly selected for district office audit, others are selected because they have a special feature that IRS is looking for, such as an unscrupulous preparer, but most are selected because the computer and/or the manual screener have determined that the return, in general, has good audit potential.

Because of the broad scope of this assignment, we concentrated on the DIF system, which accounts for most of the returns selected for audit by the district offices.

The DIF system enables IRS to evaluate every return's audit potential and effectively combines the speed of the computer with the experienced judgment of the classifier. An important aspect of this system is that it minimizes the potential for abuse in selecting returns for audit. Returns are selected not by the examiner who is responsible for auditing them but by the computer and the classifier working together.

The initial selection is made by the computer using formulas developed from TCMP data. Because all returns in a particular audit class are evaluated and scored by the same formula, there is little chance for abuse in this phase of the system. Abuse could only occur through some conspiracy to alter or bypass the scoring process.

Only a small percentage of returns, those with the highest scores, go to the second phase of the DIF system-- manual screening. The classifier is a necessary part of the system because by looking at the entire return, something the computer can't do, and by drawing on his experience he can weed out those returns that in his judgment do not warrant audit.

In making these judgmental decisions, the classifier might justifiably consider the taxpayer's place of residence and/or occupation. For example, a taxpayer might have claimed an unusually large casualty loss which contributed to the high score received by his return. The classifier, however, might notice that the taxpayer lived in an area that the classifier knew had been ravaged by floods and, on that basis, might determine that the reported loss looked reasonable and that the return did not warrant audit. Likewise, the classifier might know from past experience that persons in certain occupations tend to erroneously report certain income or expenses and his decision to select a return for audit might be based on that experience.

A classifier could abuse the system by selecting a return for audit not because of an objective determination that the return warrants audit but because he recognizes the taxpayer's name and thinks he should be audited. The chances for such abuse are minimal, however, because the classifier's decisions are subject to review by the returns program manager and, in some cases, by the group manager.

In summation, although some abuse is possible, our review of audits completed during 1973, our interviews with classifiers, and our observations of returns being classified disclosed no evidence that returns were being arbitrarily or capriciously selected for audit under the DIF system. All evidence indicated that returns selected for audit under this system were selected because, in the classifier's best judgment, they warranted audit.

In most cases, decisions to select returns for reasons other than their DIF scores are made by someone other than the person who will be auditing the return, which greatly limits the chances for abuse. There are exceptions, however. Some returns are selected by the examiner because he has determined that he needs to audit returns filed in earlier or later years by the same taxpayer or returns filed by other taxpayers that may have a bearing on the return being audited. The requisition prepared by the examiner to obtain these returns contains a code but no written explanation why the examiner needs the return and thus gives management (the group manager, the district review staff, and the internal audit staff) little basis for evaluating that need.

A question often raised about the selection process is: Why are some taxpayers' returns selected for audit year after year even though IRS never finds any major errors? Although our review disclosed no intent by IRS to harass taxpayers by repeatedly auditing their returns, we were unable, because of the wide scope of our review, to fully inquire into this matter. We are, however, currently reviewing the question of repetitive audits at the Joint Committee's request.

RECOMMENDATION TO THE COMMISSIONER  
OF INTERNAL REVENUE

To facilitate review and thus further protect against abuse, we recommend that IRS require its examiners, when requesting a return, to explain on the requisition why they need the return so that the request can be adequately evaluated.

IRS COMMENTS AND  
OUR EVALUATION

In commenting on a draft of our report by letter dated August 10, 1976 (see app. I), the Commissioner noted that:

"\* \* \* While a separate written explanation of the specific reason for the request is not required, a code number which identifies that purpose must be placed on the form by the requestor (e.g., Code 40, Prior Year Return, and Code 41, Subsequent Year Return). In most instances, the system of codes is sufficiently comprehensive to permit a supervisor or other reviewer to determine why the return was requested. \* \* \* Our procedures further require that all requisitions be approved in writing by the examiner's immediate supervisor and any questions regarding the need for securing that return would be discussed before the requisition is approved."

We do not believe that the present system of codes is sufficient to permit a supervisor or other reviewer to evaluate an examiner's request for a return. In addition to the two codes cited by the Commissioner, other codes used by examiners in requesting returns include code 50, Partner; code 53, Family Member; and code 57, Employee or Employer of Taxpayer. These codes only tell the reviewer what relationship the requested return bears to the return being audited. Without a written explanation, the reviewer has no way of knowing, by looking at the requisition, why the examiner needs the return.

Questions that go unanswered without a written explanation include: why does the examiner want the taxpayer's prior year's return or his partner's return? what did the examiner find in auditing the taxpayer's return that aroused his interest in the requested return? and how significant are the issues that the examiner wants to pursue on the requested return? In the absence of a written explanation that addresses such questions, there is little assurance that the examiner has a valid basis for his request.

We recognize that the examiner's supervisor may ask him why he is requesting the return, but there is no assurance that he will. Also, other reviewers, like the district review staff or internal audit, would be less likely to question the examiner; even if they did it would occur after the fact.

In summation, a written explanation would require the examiner to spell out his purpose in requesting the return which, in turn, might deter him from making unjustified requests and would provide an audit trail for future review.

IRS said it was making a comprehensive review of all its codes to insure that they are properly defined. We would hope that in so doing, IRS will address the concerns we have raised. Although we are recommending that examiners be required to provide written explanations in addition to the codes, we are not precluding the possibility that IRS may be able to redefine its codes so as to negate the need for such explanations in many cases.

## CHAPTER 4

### EVALUATION OF DIF EFFECTIVENESS

The Internal Revenue Service measures effectiveness of the discriminant function system by the average tax change per audit and the percentage of returns resulting in no tax change after audit (the no-change rate). IRS contends that the system becomes more effective when tax changes are higher and no-change rates are lower.

To determine DIF's effectiveness, we examined the procedures used by IRS in conducting four tests involving relationships in which DIF would be expected to prove superior if it was meeting its objectives:

- A comparison of the audit results for returns selected using DIF with the results for returns selected before DIF.
- A comparison of DIF with other computerized selection techniques and with manual selection.
- A comparison of DIF with random and perfect selection.
- A comparison of the 1973 DIF formulas with previous formulas.

We also independently evaluated DIF's effectiveness through a correlation analysis of DIF scores with the no-change rate.

The four comparisons made by IRS and our correlation analysis showed DIF to be effective.

#### COMPARISON OF DIF WITH PRE-DIF

IRS compared the audit results for returns selected using DIF with the results for returns selected before DIF was implemented.

<u>Audit class</u>	<u>Pre-DIF machine-identified returns closed in fiscal year 1969</u>		<u>DIF-scored returns closed in fiscal year 1973</u>	
	<u>Yield per return</u>	<u>No-change rate</u>	<u>Yield per return</u>	<u>No-change rate</u>
<b>Nonbusiness:</b>				
Low, standard	\$ 89	45.2	\$ 144	35.7
Low, itemized	80	47.3	182	24.4
Medium	123	47.4	269	34.3
High	1,822	42.5	5,496	31.6
<b>Business:</b>				
Low	162	43.6	762	26.3
Medium	384	37.7	1,019	26.3
High	1,450	31.8	4,033	25.3

For every audit class the DIF-scored returns produced a higher yield per return. The higher yield ranged from a low of \$55 for low nonbusiness, standard-deduction returns to a high of \$3,674 for high nonbusiness returns. Even considering the effect of inflation during the 4 years, DIF still appears more effective.

In analyzing the results of this comparison, however, we found that not all the data could be strictly interpreted to show that DIF was more effective. Audit results are influenced by factors other than the specific selection method, such as the manual screening process, the taxpayer compliance climate, and any change in IRS policies or procedures that would directly or indirectly influence the scope and quality of an audit.

Under both the DIF and pre-DIF selection methods, returns are and were manually screened by classifiers before being sent to examiners for audit. The classifiers' effectiveness or ineffectiveness in identifying returns and issues that warrant audit can affect audit results.

The taxpayer compliance climate also can affect audit results. For example, if voluntary compliance improves, there would be less tax adjustments for IRS to find. Therefore, the no-change rate could increase and yield per return could decrease even though the effectiveness of the selection procedure remained unchanged.

Any change in IRS policies and procedures affecting the scope and quality of an audit as, for example, a significant change in audit technique or average time per audit, could influence no-change rates and yield per return statistics.

COMPARING DIF WITH OTHER COMPUTERIZED SELECTION TECHNIQUES AND MANUAL SELECTION

In 1966 and 1967, IRS made a limited test to determine the feasibility of the mathematical selection of returns for audit. The test involved a sample of approximately 51,000 nonfarm business returns with adjusted gross incomes under \$10,000 filed in 1965. Three percent of these returns were selected for audit by two DIF formulas, two other computer techniques, and a manual selection method. The returns selected under each method were then thoroughly audited. Three statistical measures were used to evaluate the five selection methods: (1) average tax change per return, (2) average tax change per staff-hour of examination time, and (3) no-change rate.

<u>Statistical measure applied to test results</u>	<u>DIF formula developed by IRS</u>	<u>DIF formula developed by private contractor</u>	<u>Manual</u>	<u>Computer technique 2 (note a)</u>	<u>Computer technique 1</u>
Average tax change per return	\$376	\$342	\$354	\$294	\$142
Average tax change per staff-hour	\$ 22	\$ 20	\$ 22	\$ 19	\$ 11
No-change rate	42.6	48.7	40.9	46.7	52.3

a/Computer technique 2 was the selection method being used by IRS at the time of the test.

The two DIF formulas generally produced better results than the two other computer techniques and were competitive with the manual selection technique (a system in which all returns were screened manually and which was considered too expensive by IRS).

Because the test involved only one type of return, the results were applicable to only that type of return. Nevertheless, the test provided an indication of DIF effectiveness.

DIF COMPARED WITH RANDOM  
AND PERFECT SELECTION

IRS compared DIF results with those of random and perfect selection. A DIF score computed for each return audited under the Taxpayer Compliance Measurement Program became the basis for ranking the returns. Once the ranking was completed by audit class, the average tax change per return, at a predetermined level of audit coverage, was computed. This average tax change represented the results that would have been obtained if the returns had been selected for audit based entirely on the DIF score.

These results were compared, assuming the same level of audit coverage, with (1) the average tax change for all TCMP returns in each class--this represents random selection--and (2) the average tax change for the TCMP returns ranked by amount of tax change--this represents perfect selection.

<u>Audit class</u>	<u>Percent of audit coverage (note a)</u>	<u>Average tax change per return</u>		
		<u>DIF</u>	<u>Random selection</u>	<u>Perfect selection</u>
<b>Nonbusiness:</b>				
Low	1.3	\$ 230	\$ 36	\$ 745
Medium	1.5	802	93	1,800
High	11.7	9,460	2,178	15,312
<b>Business:</b>				
Low	2.0	940	230	3,265
Medium	2.4	2,755	345	5,512
High	12.6	5,781	1,662	10,590

a/This is the approximate percentage of returns by class that IRS audits in a fiscal year.

This test shows DIF vastly superior to random selection. The comparison to perfect selection, however, clearly indicates room for improvement. It should be noted that IRS does not select returns for audit purely on the basis of DIF scores, as was done in this test. Manual screening has always been an integral part of IRS' selection process.

1973 DIF FORMULAS COMPARED  
TO PREVIOUS FORMULAS

The DIF formulas used to score returns filed for tax years prior to 1973 were based on 1965 TCMP data. Newer DIF formulas were developed based on the 1969 TCMP.

To determine whether the new formulas were as effective as the old, IRS had a group of returns scored and ranked by both sets of formulas. Then, given a specific level of audit coverage, the results of the two rankings were compared.

<u>Audit class</u>	<u>Old formulas</u>		<u>New formulas</u>	
	<u>Average tax change per return</u>	<u>Percent no change</u>	<u>Average tax change per return</u>	<u>Percent no change</u>
Nonbusiness:				
Low	\$ 155	33.4	\$ 230	21.1
Medium	605	33.6	802	22.4
High	5,441	25.1	9,460	26.3
Business:				
Low com- mercial	665	29.1	1,035	21.6
Low farm	309	25.9	632	23.1
Medium com- mercial	1,233	13.5	3,013	13.5
Medium farm	707	19.7	1,451	12.7
High com- mercial	4,393	22.1	4,907	19.4
High farm	8,024	29.7	11,399	18.3

In terms of both average tax change and the important no-change rate, the new DIF formulas proved much more productive.

CORRELATION ANALYSIS OF DIF  
SCORES WITH NO-CHANGE RATE

The foregoing comparisons were made by IRS. We also independently analyzed the results of IRS audits to try to determine, through techniques of statistical inference, whether DIF was effective.

IRS provided us with the audit results for all returns selected under the DIF system during tax year 1973. The listing provided a unique opportunity to test IRS' contention that DIF was effective. We determined the degree of relationship between DIF scores and the no-change rate by making a

correlation analysis. Correlation analysis provides an index (correlation coefficient), or measure, of degree of relationship between two variables, in this case DIF scores and the no-change rate. The expected relationship was as follows--since the purpose of DIF is to measure the likelihood of a tax change after audit, there should be a measurable relationship between high DIF scores and low no-change rates and low DIF scores and high no-change rates. To measure this relationship, we ranked both the DIF scores and the no-change rates, with the highest DIF score ranked first and the lowest no-change rate ranked first. The details of our analysis are shown in appendix V. The results of our analysis are shown below.

<u>Audit class</u>	<u>Correlation coefficient</u>	<u>Interpretation of coefficient (note a)</u>
<b>Nonbusiness:</b>		
Low; standard	.94	Very high correlation
Low; itemized	.91	Very high correlation
Medium	.89	High correlation
High	.85	High correlation
<b>Business:</b>		
Low commercial	.41	Moderate correlation
Low farm	(b)	
Medium commercial	.70	High correlation
Medium farm	(b)	
High commercial	.82	High correlation
High farm	(b)	

a/The following interpretation was used to evaluate the degree of correlation:

Less than .20 slight; almost negligible relationship  
 .20 to .39 low correlation; definite but small relationship  
 .40 to .69 moderate correlation; substantial relationship  
 .70 to .90 high correlation; marked relationship  
 Greater than .90 very high correlation; very dependable relationship

b/The number of returns audited in this class was too small to provide a basis for analysis.

Of the seven formulas for which a sufficient number of returns were audited to provide a basis for analysis, two exhibited a very high correlation, four a high correlation, and one a moderate correlation. This indicates that the higher the DIF score on a return, the more likely it is that an audit of that return will result in a tax change. We believe that this analysis provides one of the more convincing arguments for DIF effectiveness.

#### EFFECT OF CLASSIFIERS ON DIF RESULTS

We found general agreement within IRS that the classifier influences the results obtained from DIF-selected returns because he manually screens all such returns and makes the final decision on whether they should be audited and, in most cases, decides which issues should be covered during the audit. We could not determine the extent and direction of this influence and its effect on the DIF system because any test would have required us to exercise some control over the classifiers and to disrupt IRS' operations.

#### EVALUATION OF EQUITY IN RETURN SELECTION PROCESS

Data developed from the 1969 TCMP--the last TCMP for which data on all audit classes was available--showed that 8 percent of the taxpayers overassessed their liabilities and thus were due refunds. In contrast, a review of the audits performed on DIF-selected 1969 returns revealed that only 3 percent involved overassessments.

These statistics indicate that IRS, through the DIF system, is more successful in identifying and correcting underassessments than it is in identifying and correcting overassessments. As an indication of the money involved, audits of DIF-selected 1969 returns resulted in refunds of \$8.3 million to overassessors, or about \$367 per return. If the percentage of audits involving overassessors had been 8 percent instead of 3 percent, refunds would have amounted to about \$22.1 million, assuming the same average refund per return.

The extent to which returns involving overassessments are not identified and corrected is determined primarily during the selection process rather than during the audit process. It is during the selection process that the returns to be audited are selected, and it is then that the issues to be covered in most of these audits are identified. Thus the examiner has little control over the number of audits that involve overassessments.

To determine which phase of the selection process, the computer or the classifier, is more responsible for excluding returns involving overassessments from audit consideration, we analyzed 1969 TCMP data ranked by DIF score for the medium nonbusiness class and found that the classifier is the primary contributor.

	Percentage of overassessors found after audit for medium nonbusiness class at 1970 <u>coverage levels</u>
Random selection (1969 TCMP)	11.3
1969 TCMP ranked by DIF score (excludes the classifier)	10.3
Results of audits on 1969 DIF- selected returns (includes the classifier)	5.4

IRS indicates that the objective of DIF is to discriminate between those returns with large tax change potential and those with lesser tax change potential. In developing the DIF formulas, only the characteristics of two groups--those with little or no tax change and those with a tax increase of a certain level or more--were used. IRS officials told us that, in their opinion, although the characteristics of the overassessor are not considered in the DIF formulas, he is fairly represented in the current system because IRS has found that returns involving large overassessments tend to be scored like returns involving large underassessments. A high DIF score, however, does not mean that the return will be audited. That decision is made by a classifier.

Our discussions with IRS classifiers helped to explain why overassessors were not being audited to the extent that they are found in the general population. Classifiers do not ignore the overassessor; it is just more difficult for them to identify overassessors from data on the return. A classifier, for example, might select a return for audit because the taxpayer is claiming deductions that are usually unallowable, or suspiciously rounded, or inconsistent with the taxpayer's occupation. Except in unusual cases, however, a classifier has no way of knowing whether a taxpayer might be entitled to a deduction that he failed to claim and thus would have little basis for selecting that return for audit.

Likewise, a classifier might select a return for audit if the reported income and deductions apparently leave the taxpayer with insufficient funds, considering his family size, to meet normal living expenses. Such a situation might indicate unreported income or inflated deductions. A classifier, however, would have no reason for selecting a return for audit if the reported income and deductions left the taxpayer with more than enough to meet his normal living expenses.

Classifiers told us that, when information on the return indicates a possible overassessment, it is noted and treated the same as any other potential audit item. Examples would be (1) a taxpayer including his Federal income tax refund in gross income or (2) a divorced taxpayer including child support payments in gross income.

#### TAILORING TCMP AND DIF TO LOCAL NEEDS

Under TCMP, IRS develops statistics based on a national sample of returns and, as a result, DIF formulas and compliance statistics developed from TCMP have only been used to describe national conditions.

Each IRS district performs audits in support of TCMP and forwards the data to the IRS Data Center. There it is combined with data from other districts but no district and few regional statistics are developed, even though many programs dealing with the selection of returns for audit, manpower allocation, and compliance are controlled at the local level.

Because the DIF formulas developed from TCMP data are based on national characteristics, returns filed by all taxpayers in a given class are scored by the same DIF formula no matter where the taxpayers live. The same formula is used even though it is generally believed within IRS that taxpayer characteristics differ from one section of the country to another.

In 1976, IRS initiated a study to determine whether the returns selection activity suffers in certain districts because of local factors which are not accounted for in the development of the national DIF formulas. The first step of the study involves a determination as to why some districts appear to be achieving better results under DIF than other districts. The problems in these latter districts may be due to DIF or may be due to district practices in implementing DIF. If it is determined that DIF is the root of the problem, then consideration will be given to developing a

supplementary, objective means to adapt the national DIF formulas to local conditions.

### CONCLUSIONS

IRS has developed an effective and highly sophisticated method for managing its complex audit responsibility. The effectiveness could be improved, however, if IRS made TCMP data available for use in managing local programs, measured the effect of classifiers on the DIF system, and devised a system to insure that a representative number of returns involving overassessments are audited.

We recognize that TCMP involves a national sample and that local TCMP data, because of its relatively small size, may have a larger sampling error than the national data. However, the need for local data in managing local programs demands that IRS explore ways of developing it.

We could not determine the effect of classifiers on the DIF system because any test would have required us to exercise some control over the classifiers and to disrupt IRS' normal operations. IRS, however, can and should measure this effect.

A test that IRS might consider in accomplishing this objective is one similar to that described on page 31. In such a test, IRS could

- rank TCMP-audited returns by DIF score and determine the average tax change at a predetermined level of audit coverage,
- have classifiers select for audit those TCMP returns they deem most worthy of audit, using the same predetermined level of coverage, and
- compare the average tax change that would have resulted, if the returns selected by the classifiers had been audited, with the average tax change that would have resulted, if the selection had been based solely on the DIF score.

IRS has expressed reservations about this particular test because (1) the TCMP returns would be for old years and (2) classifiers would know a test is being conducted and might not screen the returns as they usually do, thus biasing the results. We believe that (1) the TCMP returns could be classified before they are audited, rather than after, which should

resolve the age problem and (2) knowing it is a test will not significantly affect the classifiers because they will be asked to do what they always do (use their judgment to evaluate audit potential). There is little reason to believe that they would or could do a better job under test conditions. In any event, we are presenting this test only as one that IRS might consider; there may very well be a better way.

A test to determine the effect of classifiers on the DIF system might also give IRS a better indication of the extent to which returns involving overassessments are not being selected for audit. Our review indicated that, under the current selection process, taxpayers who overassess their taxes are less likely to have their returns selected for audit than are taxpayers who underassess their taxes, primarily because it is difficult for classifiers to identify overassessors.

A classifier selects for audit those returns which, in his best judgment, have good audit potential. In theory, this means not only returns involving underassessments but also returns involving overassessments. In practice, however, returns involving overassessment tend to fall by the wayside because a classifier has virtually no way of knowing, by looking at a return, that the taxpayer has overassessed himself.

IRS has two alternatives. The first would be to continue operating as it now is, relying on the classifier to select returns on the basis of judgment and experience. As such, returns involving overassessments would still, as a general rule, not be selected for audit. This alternative ignores the fact that the Government has no right to overpayments made by taxpayers faced with what is generally recognized to be a complex tax law.

The second alternative would be to devise some system that would insure that a representative number of returns involving overassessments are audited. This might require something like a secondary DIF system designed to score returns as to their potential for tax change in favor of the taxpayer. The highest scored returns could then be automatically selected for audit because no purpose would be served in having them screened by classifiers. The feasibility of any such system can be determined only by IRS.

RECOMMENDATIONS TO THE COMMISSIONER  
OF INTERNAL REVENUE

We recommend that IRS

- evaluate the feasibility of making TCMP data available for use in managing local programs,
- measure the effect of classifiers on the DIF system, including the extent to which returns involving over-assessments are not being selected for audit, and
- evaluate the feasibility of devising a system to insure that a representative number of returns involving overassessments are audited.

IRS COMMENTS AND OUR EVALUATION

The Commissioner noted that (1) current TCMP data would be of little use in managing local programs because of the high variability of the data at the district level and (2) the TCMP sample size would have to be increased substantially to acquire reliable district data. We recognize the development of statistically reliable data may require an increase in the TCMP sample size which IRS may not be able to do. We believe, however, that certain data, even if not statistically reliable, could be useful to regions and districts. For example TCMP data on the types of errors that taxpayers most often make in preparing their returns could be helpful to regions and districts in managing their classification and audit activities. IRS officials told us that such information is available, at least on a regional basis, but that it has not been routinely disseminated to the regions. We believe IRS should give more consideration to the usefulness of this and other TCMP data at the local level.

On October 19, 1976, we met with IRS officials to further discuss IRS' comments on our draft report and our reaction to these comments. We were advised, at that meeting, that IRS does use TCMP data in developing training programs for classifiers. We did not have time to verify this or to evaluate the extent to which TCMP data is used as a training aid. Training programs would appear, however, to be practical vehicles for disseminating TCMP data to the local level and such a practice would be consistent with our recommendation.

Concerning overassessors, IRS noted that in its current updating of DIF formulas for individual returns, the absolute amount of tax change (increase or decrease), rather than just

tax increase, is being used as an indicator of audit potential. IRS also agreed to explore the feasibility of developing a separate DIF system to identify returns with significant over-assessment potential if its continued monitoring of the problem discloses a systemic bias. IRS did not agree, however, that it should conduct a test to measure the effect of classifiers on the DIF system including the extent to which returns involving overassessment are being overlooked for audit. It argued that:

"To implement such a test in a time of limited research and operational resources raises a question of priorities, especially in view of the finding that 'tests conducted by IRS and GAO show this selection system (i.e., including manual screening) to be effective.' But more importantly, the Service plans to revise its instructions for screening high-scored DIF returns so that the number of returns that are accepted as filed will be reduced. In general, under this revision, high-score returns would be selected for audit unless sufficient data has been submitted as a part of the return to support questionable items. Since returns with significant potential over-assessments also tend to have high scores, the emphasis on selecting high-score returns, unless there is apparent justification to the contrary, would help reduce screening bias towards selection of underassessed returns."

We discussed the proposed revised instructions referred to in IRS' comments with officials of the Audit Division. As we understand it, the instructions will significantly alter the classifier's role in the DIF system. In the past, classifiers would screen high DIF-scored returns and select those that, in their judgment, warranted audit. The proposed instructions will tell classifiers, in effect, that the high DIF-scored returns have already been selected for audit by the computer and that their job is to weed out those returns that do not warrant audit either because the taxpayer has attached something to his return explaining an otherwise questionable item or because the classifier is aware of a local condition which would explain an unusually large deduction (such as a natural disaster causing a large casualty loss). These are things that the computer cannot consider in scoring the return. The classifier would no longer be able to weed out a return as not warranting audit simply because the return in general appeared to have no audit potential.

The changes being made in updating the DIF formulas and the proposed revised instructions for classifiers should, if effectively implemented, better insure that a representative number of returns involving overassessments are audited. We do not believe, however, that the revised instructions will negate the need to measure the classifier's effect on the DIF system.

The classifier serves a dual role in the DIF system--he selects the returns to be audited and, in most cases, decides which issues should be covered during the audit. The proposed instructions would serve to restrict the first of these two roles but would not affect the second. The classifier's success in pinpointing the issues that should be audited has significant impact on DIF. For example, if the classifier overlooks major issues, the audit will result in a lesser tax change and maybe even no tax change--thus diminishing the effectiveness of DIF. We believe that IRS should measure the extent and direction of the classifier's impact.

Although, as IRS noted, we did find the DIF system effective, we did not mean to imply that it could not be improved. The system involves two stages--the computer and the classifier. Tests conducted to date have proven rather conclusively that the computer phase is effective, but little has been done to evaluate the effectiveness of the classifier.

The question of priorities raised by IRS is a valid one. As already indicated, however, we believe that the classification activity is more important than IRS seems willing to admit. In our opinion, IRS should be concerned whether classifiers are selecting the best issues for audit and, whether as a result, their impact on the DIF system is beneficial. Pinpointing classifiers' problems in selecting the best issues for audit could be instrumental in improving the classification function and reducing the no-change rate--one of the main reasons for establishing the DIF system and a continuing IRS goal. We suggested a controlled test that might be used to achieve this objective, but there may be less costly options.

At our October 19 meeting, IRS appeared to recognize the validity of measuring the effect of classifiers on the DIF system and said it would consider ways to do it.

## CHAPTER 5

### HOW IRS DETERMINES THE NUMBER AND TYPES OF RETURNS TO BE AUDITED--LONG-RANGE STRATEGY

The primary objective in selecting tax returns for audit is to identify those returns with the highest potential for tax change within an audit class and, thereby, effect a high degree of voluntary taxpayer compliance. Before it can meet this objective, however, the Internal Revenue Service must know what the voluntary compliance levels are and must determine the audit effort it wants to direct toward each audit class to improve or maintain those levels. The first step in this process is the development of a long-range audit strategy. The second step--development of an annual plan--is discussed in chapter 6.

#### VOLUNTARY COMPLIANCE

The present system of income taxation depends on voluntary compliance; that is, the willingness and ability of taxpayers to assess their taxes correctly. The data IRS uses to determine voluntary compliance levels is obtained from specialized audits of tax returns randomly selected under the Taxpayer Compliance Measurement Program. In 1964 the individual return phase of this program was begun. A random sample of 94,000 individual 1963 income tax returns was assigned to experienced examiners for audit. The examiners recorded detailed information about each return as originally filed and as corrected after examination of the taxpayers' books and records and other records available to IRS and after interviews with the taxpayers and their counsels. A tape file was then compiled, containing the information from the returns as filed and from the audits.

TCMP audits, unlike other audits, involve a thorough review of the entire return. Examiners are instructed to review every item, regardless of the dollar amount, and to be especially alert to discovering unreported income and to allowing taxpayers additional deductions or credits to which they are entitled.

The TCMP audits of 1963 returns and those of 1965, 1969, and 1971 returns have served as a basis for many of the programs and much of the planning within IRS. Audit work began in 1974 on 51,000 randomly selected 1973 returns and a sample of about 4,000 returns filed by taxpayers who had been audited during the 1969 TCMP cycle. The purpose of the second sample was to help assess the impact of audits on compliance.

### Voluntary compliance rates

Using TCMP data, IRS measures voluntary compliance by relating the tax liability voluntarily reported by persons who have filed returns to the total tax liability those same persons should have reported. <sup>1/</sup> These measurements show voluntary compliance on the decline.

<u>Audit class</u>	<u>Percent voluntary compliance</u>			
	<u>1965</u>	<u>1969</u>	<u>1971</u> <u>(note a)</u>	<u>1973</u>
Low nonbusiness:				
Standard	95.4	95.2	-	93.7
Itemized	91.9	88.5	86.0	85.3
Medium nonbusiness	96.6	96.1	95.9	95.7
High nonbusiness	95.8	94.1	-	95.2
Low business	78.0	68.7	63.5	56.6
Medium business	90.7	87.8	-	86.0
High business	<u>93.3</u>	<u>91.2</u>	-	<u>90.6</u>
Total	93.8	92.7		92.3

a/Only three classes were covered in the 1971 TCMP.

As can be seen, total compliance dropped from 93.8 percent in 1965, to 92.3 percent in 1973. At the same time, audit coverage had dropped from 4.7 to 3.0 percent. Data obtained during the 1969 TCMP cycle showed that, due in part to reduced compliance, unreported tax liability (underreported liability less overreported liability) had increased

---

<sup>1/</sup>IRS recently recomputed its voluntary compliance rates. In the original computations, the tax that should have been reported was determined by adding tax overpayments and underpayments to the tax voluntarily reported. In the revised computations, only underpayments were added to the amount voluntarily reported. The compliance rates cited in this chapter are the original computations because those are the ones IRS was using at the time of our review. The recomputations resulted in increases to the rates shown in this chapter of not more than 1.7 percent and have no effect on the matters discussed.

by \$3.2 billion--the difference between the \$1.8 billion tax gap <sup>1/</sup> in 1965 and the \$5.0 billion tax gap in 1969. IRS' most recent statistics, based on 1973 TCMP data, showed the tax gap to be \$6.7 billion.

#### Forms of noncompliance

Noncompliance takes several forms, including (1) improperly claiming credits, deductions, and exemptions, (2) failing to report income, and (3) failing to file a tax return. The full extent of noncompliance is not known because IRS has not fully measured unreported income and because IRS does not consider nonfilers in measuring compliance.

Although examiners assigned TCMP returns are instructed to be alert to unreported income, this form of noncompliance is more difficult to identify than incorrectly reported deductions, credits, and exemptions. IRS can require support from the taxpayer for the items included in the tax return. It is much more difficult, however, to prove unreported income because a taxpayer will be reluctant to voluntarily reveal this information.

Compliance is measured by reference to the tax liability reported and unreported by persons who have filed returns. It does not consider the tax liability that should have been reported by persons who did not file returns. The exclusion of nonfilers in measuring compliance does not affect the planning process. Since the objective of the planning process is to direct audit resources where needed, IRS, in planning its audit activity, is concerned only with the compliance associated with filed returns because those are the only returns that can be audited. Also, although nonfilers are excluded from IRS' measurement of compliance, they are not ignored by IRS. Responsibility for securing delinquent returns and tax payments from nonfilers is a function of IRS' collection activity rather than its audit activity which is the subject of this report. In line with this responsibility, the Collection Division does have enforcement programs directed at nonfilers.

#### Comparison of voluntary compliance and the no-change rate

IRS measures voluntary compliance by comparing tax dollars voluntarily reported with tax dollars that should

---

<sup>1/</sup>Total unreported tax liability less additional tax liability disclosed through audits.

have been reported. Such a measure does not provide any indication of the number of persons who correctly report their tax liability--the no-change rate. The no-change rate, which is also developed from TCMP data, is the ratio of the number of audits that do not result in a change in tax liability to the total number of audits.

The level of voluntary compliance reported by IRS differs greatly from that indicated by the no-change rate. For example, projections based on 1969 TCMP results show that 68 percent of the taxpayers (50.5 million) paid their proper taxes, <sup>1/</sup> while IRS reported, for the same year, that the voluntary compliance level for all taxpayers filing individual returns was 92.7 percent. A comparison by audit class of IRS estimates of voluntary compliance and the percentage of taxpayers who correctly reported their taxes within \$25 (the no-change rate) for 1969 is shown below.

<u>Audit class</u>	Percent no change	Percent voluntary compliance
Nonbusiness:		
Low, standard	92	95.2
Low, itemized	62	88.5
Medium	55	96.1
High	39	94.1
Business:		
Low	45	68.7
Medium	38	87.8
High	30	91.2

Although the rates differ, we believe, from the standpoint of equity between audit classes, that IRS is correct in developing its audit strategy on the basis of dollar compliance instead of no-change rates. Consider, for example, two audit classes: In one class, 45 percent of the taxpayers are correctly reporting their taxes but the class, as a whole, is only reporting 69 percent of the taxes it should be; while in the other class, only 30 percent of the taxpayers are correctly reporting their taxes but the class, as a whole, is

---

<sup>1/</sup>The 68 percent includes about 17 percent who were within \$25 of paying their correct taxes. We included these 17 percent because, in practice, when an examiner audits a return and finds a tax deficiency of only a few dollars, he usually closes the case "no-change" instead of assessing the taxpayer.

reporting 91 percent of the taxes it should be. IRS would appear correct in expanding its audit effort against the first class rather than the second because, relatively speaking, the noncompliers in the first class are paying a lower percentage of what they should than are the noncompliers in the second class.

HOW IRS DETERMINES ITS  
LONG-RANGE AUDIT STRATEGY

IRS has attributed the continuing decline in compliance, in large part, to the fact that increases in audit staff and in the number of audits have not kept pace with the increase in the number of returns filed. Staffing has been affected by hiring constraints and special programs, such as when agents and auditors were assigned to monitor compliance with the Economic Stabilization Program. Because IRS sees the drop in compliance as the effect of reduced audit effort, it has advocated a strong audit program. It believes that such a program will encourage compliance with the tax law and help assure the public that each taxpayer is paying his fair share of the tax burden.

To facilitate discussion of how IRS determines its long-range audit strategy, we will explain the procedures followed by IRS in developing its strategy for the 5 years ending with fiscal year 1979. IRS' long-range planning covers corporate as well as individual returns. Our discussion, however, will be limited to individual returns.

When IRS began developing its strategy, voluntary compliance rates were on the decline; the rate for the low business class had dropped below 70 percent; and the rates for the low nonbusiness (itemized) and the medium business classes had dropped below 90 percent.

With this in mind, IRS considered three basic audit strategies, each of which was designed to emphasize slightly different objectives.

1. The compliance-oriented strategy.
2. The maximization of direct yield strategy.
3. The balanced strategy.

The compliance-oriented strategy was designed to improve compliance in audit classes where it was considered low, while at the same time maintaining compliance in other audit classes.

The maximization of direct yield strategy was designed to achieve just that. Staff was allocated first to the audit class with the highest yield-to-cost ratio. <sup>1/</sup> As the more productive returns in that class were audited, the ratio would decline until it dropped below that of one of the other classes. At that point, additional staff would be allocated to the new class. The process continued until all audit staff were assigned and the yield-to-cost ratios were equal for all classes. This strategy, if adopted, would result in the highest possible return per dollar spent on audit. Such a strategy, however, could seriously affect compliance in the low-yield classes because audits would not be planned for those classes until the yield-to-cost ratios of the other classes fell below the ratios of the low-yield classes.

The balanced strategy was a mix of the compliance-oriented and maximization of direct yield strategies. It was designed to improve compliance in the audit classes where compliance was considered low and to allocate remaining resources to the rest of the audit classes on the basis of yield.

Each strategy was evaluated in terms of estimated compliance and potential revenues using the following mathematical relationships.

--Percent audit coverage to voluntary compliance.

--Percent audit coverage to average yield.

#### Relationship of audit coverage to compliance

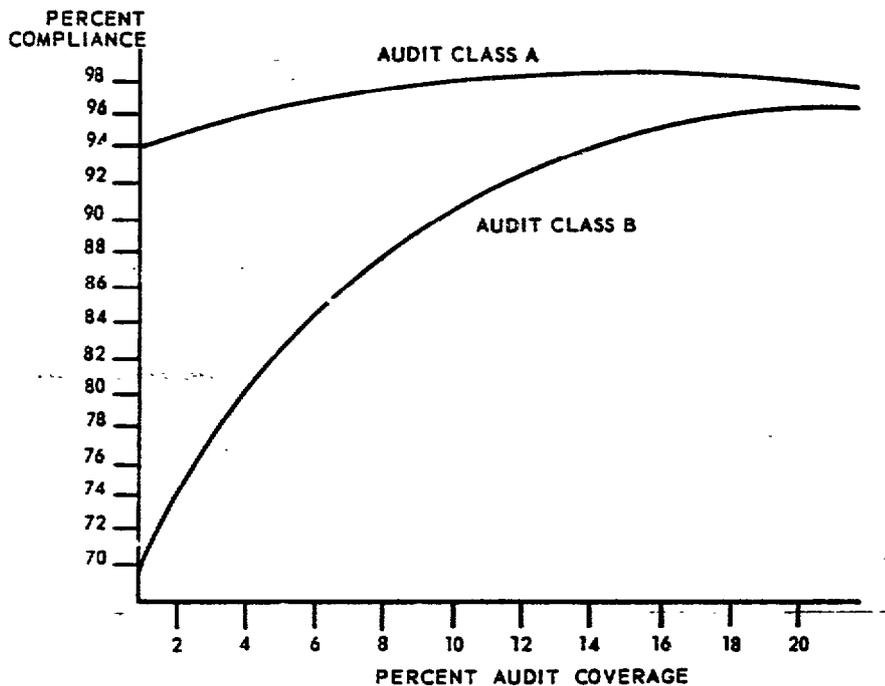
IPS believes that TCMP audits of individual returns reveal a close relationship between prior year audit coverage and voluntary compliance levels in the following year. IRS feels this relationship can best be described quantitatively by a mathematical function that represents compliance as increasing, though at a decreasing rate, as audit coverage increases. The conventional method of showing this relationship through its possible ranges is a curve.

---

<sup>1/</sup>The ratio of the additional taxes assessed as the result of audit to the cost of performing the audit.

The curve selected by IRS is known as the Gompertz curve. IRS believes this curve adequately depicts the relationship between compliance and coverage, which it sees as being such that most taxpayers will tend to improve compliance as the level of audit coverage increases although it recognizes that other taxpayers will be reluctant to change their behavior even if a relatively high probability of audit exists. IRS has developed a curve for each audit class using compliance data obtained from the 1963, 1965, 1969, and 1971 TCMPs. 1/

Once the curves have been derived, it is relatively simple to pick the compliance level applicable to a specific audit coverage and to compute changes in voluntary compliance that will result from changes in coverage. For example, the following curves show the relationships developed for two audit classes. 2/



1/At the time of our review, the curves had not yet been adjusted to reflect data obtained from the 1973 TCMP.

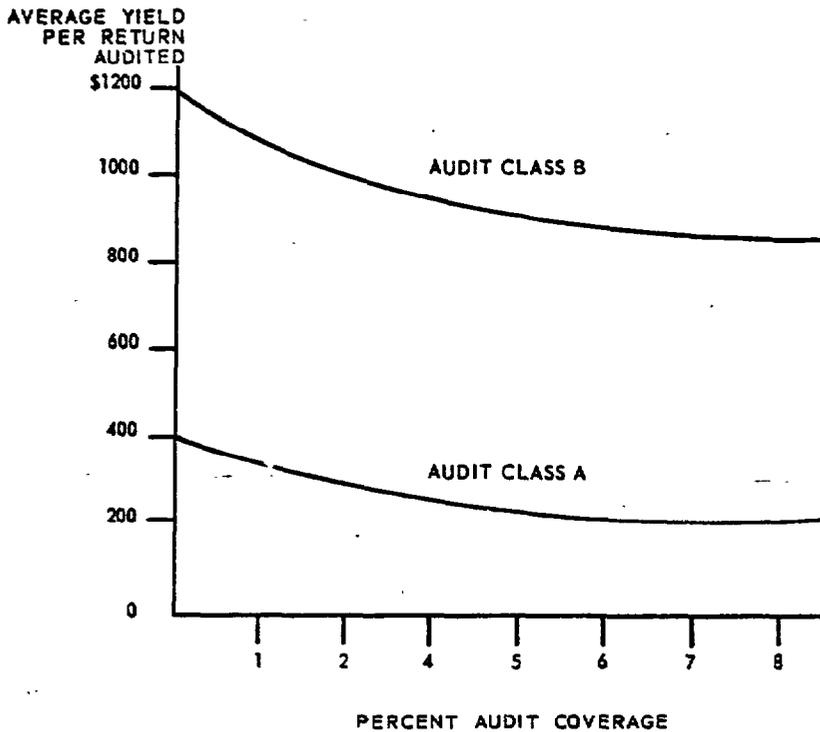
2/These and other curves shown in this chapter are intended to approximate, not mirror, the curves used by IRS.

The usefulness of these curves in the planning process can be demonstrated as follows. Assume that audit coverage is 2 percent for both classes and that an increase in personnel would allow the coverage of both classes to be raised to 4 percent or the coverage of either class to be raised to 6 percent. If the goal is to increase overall compliance, then the resources would be assigned to audit class B because such a decision would increase compliance in that class about 14 percent.

Relationship of audit coverage to yield

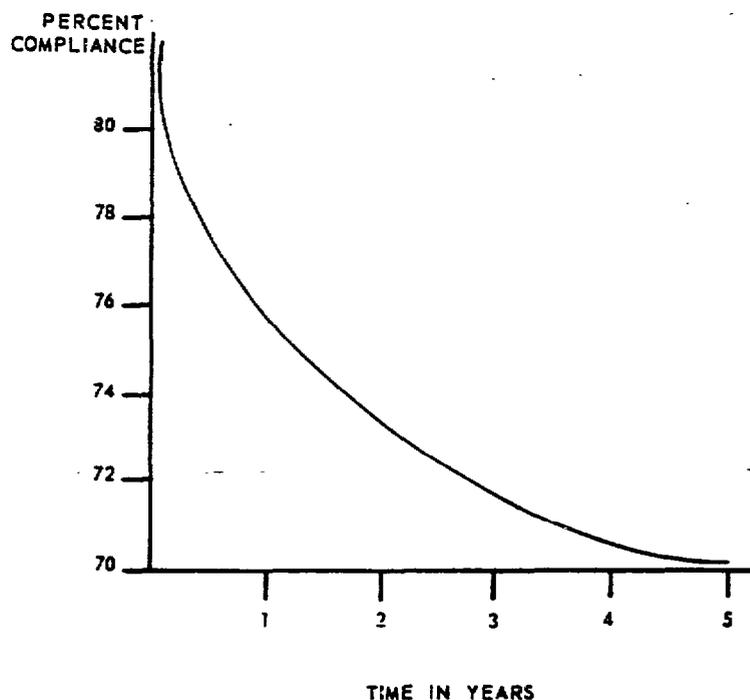
IRS has also developed direct yield curves by DIF-scoring TCMP returns and then determining the average tax change of the returns ranked by DIF scores. This change is then related to the level of audit coverage. The level of coverage and the estimated average tax change are then plotted and a relationship between coverage and yield is derived.

For example, IRS has developed the following relationship between audit coverage and average yield per return for two audit classes. By using similar curves developed for all classes of returns, IRS can project the expected yield from various levels of audit.



As noted, the relationship between coverage and yield is derived from TCMP audit results. TCMP audits, however, differ significantly from regular audits in that they are done in more depth, a higher degree of consistency is maintained, and the returns audited are randomly selected. As a result, TCMP yield data is not representative of actual operating experience. Recognizing this, IRS adjusts the TCMP-based yield relationships, using the results of regular audits, so that its estimates of direct yield will more accurately reflect actual experience.

IRS would like to include the residual effects of audit in its planning process, but until recently it had no data with which it could measure these effects. The TCMP sample of 1971 individual tax returns, however, included 3,000 returns filed by people who had been audited in the prior TCMP cycle. From this sample, IRS hoped to develop information which would allow a better understanding of the impact of an audit. IRS expected the information to show that the longer the period since audit, the less the impact on the current return. The following diagram presents a generalized form of such a curve, in terms of compliance and time.



### Results of evaluation

In evaluating the three strategies, IRS started with the compliance-oriented strategy and set desirable compliance levels for each class of return. These compliance levels were transformed into audit coverage using the relationship between coverage and compliance. The coverage was converted to staff, and IRS determined whether the staff could be recruited and absorbed into the system. After the availability of staff was established, the direct yield from audit was determined by relating coverage to yield. The other two strategies were then evaluated using the staffing computed for the compliance strategy as a base.

After evaluating each strategy's effect on compliance and yield for individual and corporate returns, IRS selected the balanced strategy because, for individual returns, it combined high yield with the highest overall compliance level.

### OPPORTUNITIES TO IMPROVE THE PLANNING PROCESS

IRS' planning process provides a sound framework for resource allocation. The process could be improved, however, if IRS (1) intensified its research into what factors, other than audit, may affect compliance, (2) included the impact of service center audits on compliance, and (3) automated the process.

### Other factors influencing voluntary compliance

According to IRS, each of its many functions, such as taxpayer assistance, intelligence, collection, and audit, contributes to the goal of improving voluntary compliance. Although there is no specific data as to which function contributes the most, IRS has traditionally considered audit coverage to be a significant, if not the most significant, factor. To support this contention, IRS points to the fact that, as the percentage of returns audited has decreased, so has the level of voluntary compliance.

Although audit coverage and taxpayer fear of possible audit may influence compliance, IRS needs to determine the influence of other factors. We believe information of this nature is a necessary prerequisite to efficient allocation of audit resources. IRS' single research effort in this area was started in 1970 and, in spite of some rather startling results, has progressed slowly since then.

The basic approach of this research was to first develop voluntary compliance measures by seven basic audit classes for several geographical areas in the United States and then attempt to explain compliance levels by factors called independent variables. Two groups of variables were considered (1) 16 internal variables which, according to the study group, represented policy factors which IRS could control, such as audit coverage, collection activity, and return complexity, and (2) 17 external variables representing social and economic factors which, the study group felt, IRS could not control, such as housing patterns, racial mixes, and employment characteristics. Using multiple regression, IRS tried to determine which of these variables best explained voluntary compliance levels.

Preliminary findings indicated that some external variables may have a significant influence on taxpayer compliance. A ranking of the 33 variables considered showed that in 6 of the 7 audit classes, the variable that most affected compliance was an external one. Furthermore, none of the internal variables were ranked in the top 10 for all 7 classes, whereas 2 of the external variables were.

Although this approach is too preliminary to be used in its present state, its potential usefulness was described in the IRS report on the project.

"Knowledge of this sort helps pinpoint not only tax administration problems, but also basic social, economic, and political problems. This project, with its broad criminal--sociological--economic approach, should throw some light on whether the traditional enforcement approach or a more 'enlightened' taxpayer service and education approach would be more effective in solving such basic problems as how to raise voluntary compliance in ghetto areas." (Underscoring supplied.)

IRS is also conducting research to determine the deterrent effect of the audit program on taxpayers actually audited. The effect manifests itself by the taxpayer correctly reporting his taxes in the future as a result of a past audit. Research on the results of regular audits shows that audits do provide a deterrent and that taxpayers are more apt to correctly report their taxes in the years following an audit. IRS is also studying the effect of audit on compliance using RCMP audits. Neither study considers the impact of audit on taxpayers not actually audited--the so-called "audit presence" or "ripple" effect. Therefore, the results of this

research apply to only about 2 percent of the taxpayers (those actually audited in a given year).

There are some indications that, because taxpayers fear the punishment the Government can administer, they are reluctant to cheat. A survey conducted for IRS in 1966 showed that 70 percent of the taxpayer respondents felt that some taxpayers were tempted to cheat but decided not to because they had heard about how the Government punishes tax cheaters through legal sanctions.

A different aspect of the question of sanctions on taxpayer compliance was addressed in a study entitled "On Legal Sanctions" by Richard D. Schwartz and Sonya Orleans published in 1967. <sup>1/</sup> The results of this study suggest that appeals to taxpayer conscience are more effective than threats of sanction, though both have some effect on taxpayer compliance. The study concludes that much work is needed to build an adequate theory of tax compliance.

Although these studies indicate possible relationships between compliance and the fear of audit, much needs to be done to further definitize this relationship and to determine other factors that affect compliance.

Impact of service center audits  
not considered in planning process

In evaluating its planning strategies, IRS sets desirable compliance levels and then determines the audit coverage needed to achieve those levels by using curves that relate audit coverage to compliance. At the time of our review, these curves related compliance data obtained from TCMP audits of 1963, 1965, 1969, and 1971 tax returns with prior years' audit coverages. In those years, all audits were done by district offices and so the audit coverage figures used in plotting the curves represented IRS' total audit effort.

Since 1972, however, many audits have been done by service centers. Because most of these audits do not fall within IRS' more restrictive definition of "audit," their impact on compliance has been virtually ignored by IRS in its planning. For example, if IRS decides that it wants to achieve a compliance level of 85 percent in the low business class, it

---

<sup>1/</sup>The University of Chicago Law Review; Volume 34, number 2; 1967.

refers to the appropriate curve to determine the audit coverage needed to meet that goal. Assuming that the curve shows 4 percent coverage is needed and that 4 percent coverage represents 150,000 audits, IRS will develop a plan calling for 150,000 district office audits of low business returns. IRS ignores the fact that many low business returns will be audited by the service centers. Although service center audits are conducted by mail, there is no reason to believe their impact on compliance differs substantially from the impact of district office audits, especially since many district audits are also handled by mail. The following table demonstrates the significance of service center audits.

<u>Audit performed by</u>	<u>Number of audits</u>	
	<u>Fiscal year 1974</u>	<u>Fiscal year 1975</u>
Service centers	712,000	1,322,000
District offices	<u>1,687,000</u>	<u>1,839,000</u>
Total	<u>2,399,000</u>	<u>3,161,000</u>

IRS has recently begun considering the impact of service center audits done under the DIF correspondence program and told us that, for the 1977 edition of the long-range plan, it intended to consider the impact of audits done under three other service center programs--claims, Federal-State cooperative audit, and social security referral. According to IRS, the examinations done under these four programs fall within its definition of "audit" because they involve a review of taxpayer records. These four programs, however, involved only 102,500 audits in fiscal year 1975--8 percent of all service center audits that year.

By omitting the impact of all service center audits in the planning process, IRS is overestimating the district office resources required to meet its long-range compliance goals for individual returns. If all service center audits were included, fewer district office audits, and thus fewer district office examiners, would be needed to attain those goals. IRS could then either decrease the number of examiners it planned to hire or redirect its resources to other types of returns such as corporate, estate, gift, excise, or employment.

In other words, if IRS determines that it needs to audit 150,000 low business returns to achieve 85 percent compliance and if it knows from past experience that the service centers may audit about 50,000 low business returns, then it should

plan for only 100,000 district office audits and determine its staffing needs accordingly.

#### Need for computerizing the planning process

IRS evaluates its strategies in terms of expected compliance and yield and then selects one of them for implementation. Its evaluation has been basically a manual one which has limited the number of strategies that could be considered. We discussed with IRS officials how computerizing the process would allow for the evaluation of more strategies and the introduction of a wider range of factors into each strategy.

IRS has since begun automating its planning process.

#### CONCLUSIONS

Overall, the method used by IRS to develop its long-range strategy represents an imaginative approach for dealing with this complex matter. IRS has expended considerable effort and used sophisticated analytical techniques to aid it in identifying problem areas, developing alternative solutions to these problems, and determining what it considers to be the "best" solution, given existing constraints. The process can be improved, however.

Because IRS' main concern in recent years has been the general decline in voluntary compliance and because it does not know exactly how audit coverage and other factors affect compliance, it should emphasize research in this area. Many questions need to be resolved.

- The deterrent effect of an audit on future compliance.
- The extent, if any, of the so called "ripple" effect.
- The effect of approaches other than enforcement on compliance.

IRS should include the impact of service center audits in measuring the relationship between audit coverage and compliance and in establishing resource requirements. In our opinion, these audits do affect compliance because, like audits done by the district office, they involve IRS contacting taxpayers about problems on their returns. In that respect, if an audit initiated by a service center is transferred to a district office, at the taxpayer's request for example,

it is considered a district office audit and thus is considered to have an impact on compliance. If the audit is initiated and closed by the service center it is considered to have no impact on compliance. We do not believe that an audit's impact will vary simply because the taxpayer's contact is with a service center rather than a district office.

Finally, by computerizing the planning process IRS could increase the number of options that can be considered during the process.

#### RECOMMENDATIONS TO THE COMMISSIONER OF INTERNAL REVENUE

We recommend that IRS:

- Expand and accelerate its research into factors which influence compliance, including those external factors identified in its previous study.
- Modify its planning process to include the impact of service center audits.
- Computerize its planning process.

#### IRS COMMENTS AND OUR EVALUATION

IRS plans to continue searching for economical ways to assess internal and external factors influencing voluntary compliance and will not hesitate to seek additional funds for this purpose as promising methodologies are identified.

IRS told us that it had executed a contract for an outside contractor to develop a computer program that will enable a rapid evaluation of multiple alternative planning strategies. IRS expected the program to be available for use in developing the 1977 edition of its long-range plan.

IRS did not agree that it should include the impact of all service center audits on compliance because

- most of these audits, such as those done under the unallowable items and head of household programs, do not require examination of the taxpayer's books and records and thus do not fall within IRS' definition of "audit,"
- it is difficult to estimate the workload by audit class that these audits would generate in any given year, and

--IRS doubts that these audits, in total, have the same overall effect on compliance as do regular audits.

We believe that IRS is being too restrictive in defining "audit" for planning purposes. From a practical standpoint, the average taxpayer would, in our opinion, consider himself audited if he were to receive a letter from IRS saying that a review of his return indicates that he has claimed an unallowable item or has erroneously used the head-of-household tax rate and that he can either justify what he did or pay the additional tax.

We believe that IRS is seeking unnecessary preciseness in its planning process by claiming that it would be difficult to estimate the workload that service center audits would generate. True, it would be difficult to estimate the precise workload but we do not expect precision. We look upon IRS' planning process as providing direction for the audit program, and, in our opinion, that direction could be improved if a reasonable estimate of the impact of service center audits were included.

IRS expressed doubt that service center audits, in total, have the same overall effect on compliance as do regular audits. There is no hard evidence to support this doubt, however, and, in fact, IRS has acknowledged that it needs to know more about the factors that affect compliance. Although service center audits may not have the same effect as regular audits on compliance, there is ample reason to believe that their effect is substantial.

IRS measures compliance in terms of dollars; the more a taxpayer underreports his tax liability, the lower the compliance rate. If a taxpayer underreports his tax liability by \$100 because of an unallowable item, for example, the effect on the compliance rate is the same as it would have been if the underreporting had been caused by the taxpayer claiming travel expenses that he could not adequately support. In truth, the average service center audit involves less additional tax than the average district office audit so that the effect of a service center audit on compliance could be considered less than the effect of a district office audit. On the other hand, the service center audits in question are, in IRS' words, "designed to correct specific items which can be readily identified as being erroneous," while many district office audits involve expenses claimed by the taxpayer for which he has inadequate support--an issue which is not always cut and dried but often a matter of judgment. Thus,

a taxpayer who has had his return adjusted by the service center for an unallowable item or the like might find it easier to understand his error and comply in the future than would a taxpayer who had had his return adjusted by the district office because of inadequate support. In this respect, the effect on compliance of service center audits might even be greater than that of district office audits.

On October 19 and 20, 1976, we talked to IRS officials about service center audits. They noted that the recently enacted Tax Reform Act of 1976 (90 Stat. 1703) prescribed different procedures for handling unallowable items. The act refers to these items as clerical errors and calls for them to be handled the same as mathematical errors. IRS wondered whether this revision would make our recommendation moot. They also restated their position that it is difficult to estimate the workload that service center audit programs would generate in any given year.

We pointed out that the change in the tax law only relates to unallowable items and that the service centers will still be contacting taxpayers about problems on their returns under other audit programs. Also we recognized the difficulty in precisely estimating service center audit workload but expressed the belief that such precision is unnecessary for planning purposes.

IRS acknowledged that service center audits do have some impact on compliance but expressed uncertainty as to the extent of that impact. It agreed to consider that question in conjunction with any assessment it makes of the internal and external factors affecting voluntary compliance. Based on its findings, it will then determine the extent to which it should modify its long-range planning process to include the impact of service center audits on compliance.

NEED FOR IRS TO PROVIDE MORE INFORMATION  
IN SUPPORT OF ITS APPROPRIATION REQUESTS

Each of the planning strategies IRS considered in developing its fiscal year 1979 long-range plan called for different compliance levels for different audit classes. The balanced strategy, for example, called for compliance levels ranging from 85 percent for the low business class to 96.7 percent for the high nonbusiness class. Thus, IRS is saying that it expects certain classes of taxpayers to meet a higher standard of compliance than other classes.

These varying compliance standards arise because IRS is not only interested in compliance but also in yield. If IRS were interested only in compliance, it would pursue a strategy that calls for all classes to meet the same compliance standard.

The question as to whether IRS should continue following a balanced strategy that calls for differing compliance levels or whether it would be practical to develop some other strategy is one that should be considered by the Congress in acting on IRS' appropriation request for audit manpower. IRS, however, does not provide the Congress with the information necessary to address that question.

In the justification for its fiscal year 1976 appropriation request, for example, IRS stated that its fiscal year 1976 audit program was based on a long range strategy which would raise the overall level of voluntary compliance, with particular emphasis on those audit classes which have a relatively low level of compliance. The justification did not include information on the specific basis for the request; namely, the audit strategies considered by IRS, the strategy selected and why, and the actual and anticipated compliance levels.

If the Congress knew the basis for IRS' request and the long-range compliance levels IRS wanted to achieve, it would be better able to evaluate that request and decide whether it wanted to commit the resources necessary to achieve those compliance goals or whether it wanted the goals revised.

#### IRS comments and our evaluation

We had recommended that IRS, in its appropriation requests, state the compliance levels it desires to achieve and directly relate these objectives to the requested appropriation. IRS took exception to this proposed recommendation by noting that it (1) does not have the necessary data to reliably predict the resultant levels of compliance for all types and classes of returns examined, such as corporate, gift, and excise tax returns, and (2) is unable to measure compliance within a fiscal year.

IRS apparently misinterpreted our recommendation. First, we were directing our recommendation only at individual returns since that is the subject of this report and IRS does have the necessary data to reliably predict compliance levels for individual returns. Second, we did not intend, by our recommendation, to suggest that IRS measure compliance within a fiscal year. What we intended to suggest was that IRS

give the Congress the benefit of the planning and thinking behind its appropriation request, as it applies to individual returns. In other words, IRS should let the Congress know what long-range planning strategies it considered, why it selected the strategy it did, what the most recent TCMP-based compliance levels are, what levels it hopes to achieve in the long run as a result of the selected strategy, and how the current appropriation request relates to that long-range plan.

As an example of the type of information that we think would be helpful to the Congress in understanding and evaluating IRS' appropriation request, IRS, in commenting on the fact that the balanced strategy called for differing compliance levels, told us that

"\* \* \* because of resource constraints, IRS is forced to accept differing levels of compliance among [audit] classes. Moreover, IRS believes that in some low compliance classes, other means such as taxpayer assistance, education and corrective legislation could prove more cost effective in improving compliance than large increases in audit coverages."

The recommendation in our draft was addressed to IRS. However, because we believe that more detailed information would assist the Congress in evaluating IRS' appropriation request and because the Congress must ultimately decide whether it needs and wants that information, we are now addressing our recommendation to the Congress.

#### Recommendation to the Congress

We recommend that the Congress request IRS to provide, as part of its appropriation request, detailed information on the planning and thinking behind its request, such as the various strategies considered, the strategy selected, and the long-range goals associated with the strategy.

## CHAPTER 6

### HOW IRS DETERMINES THE NUMBER AND TYPES OF RETURNS

#### TO BE AUDITED BY DISTRICT OFFICES--ANNUAL PLAN

With input from the national, regional, and district offices, the Internal Revenue Service develops plans for the coming fiscal year--the first step toward meeting its long-range goals.

The ultimate goal is to have a viable plan for each district while keeping the total of the 58 district plans as closely aligned with the original national plan as possible.

#### NATIONAL OFFICE

Annually, the national office prepares two work plans--a "base" plan, which assumes a carryover of current staffing levels, and an "expansion" plan, which assumes congressional approval of additional audit staff. Both plans show by audit class the number of returns expected to be filed, the percent audit coverage, the rates of examination (the number of returns that can be examined in a direct examination staff-year <sup>1/</sup>) for both revenue agents and tax auditors, the number of returns to be examined by agents and auditors, and the number of agent and auditor direct examination staff-years needed. The plan does not refer to dollars, just numbers of audits and staff-years.

The national office allocates a portion of its plan to each of the seven regions, as illustrated by the following example.

Assuming that its plan calls for the audit of 100,000 medium nonbusiness returns, the national office, by reference to DIF scores for returns filed in the prior year, determines the score to which it must go to get the returns needed--referred to as the cutoff score. If the cutoff score is 250, the national office, by reference to regional DIF scores for the prior year, determines the number of returns in each region with scores of 250 or above.

---

<sup>1/</sup>A direct examination staff-year indicates the time spent actually examining returns. The remaining time is spent on activities, such as leave, training, or taxpayer service.

The plan is then allocated to each region in the same proportion that the number of returns in the region at or above the cutoff score bears to the number of returns nationwide at or above that score.

By using DIF scores, which are indicators of audit potential and taxpayer compliance, the national office can identify those areas where there is apparently a greater degree of noncompliance and where a greater percentage of taxpayers would have to be audited to improve compliance. In other words, if two regions have the same number of returns filed and if one region has 20 percent of all the returns above the cutoff score while the other region has 10 percent, it would indicate that compliance in the first region is lower than in the second. By allocating on the basis of DIF scores, the national office provides for more audits in the region with the apparently lower compliance. This initial allocation is referred to as the optimal base plan.

After the optimal base plan is developed, the national office determines whether each region has sufficient staff to meet its plan. If a region has more or less staff than needed, the national office revises the plan accordingly.

The following table shows for each region the difference between the number of audits to be done in fiscal year 1975 under the revised base plan and the number that should have been under an optimal staffing allocation (optimal base plan).

Region	Audit class	Nonbusiness:		Low, standard		Low, itemized		Medium		High		Business:		Low		Medium		High		Total																				
		Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax	Revenue tax																					
North Atlantic	Revenue tax	1,432	-	(435)	(14)	(26,217)	(5)	6,125	28	3,810	23	737	(4)	661	606	3	349	461	661	13,847	2,995	23,673	(2,040)	(2,629)	(585)	2,389	13	13	12,225	1,295	11,187	(1,888)	(92,875)							
	Revenue tax	-	-	(14)	(1)	(206)	(2)	(349)	(10)	(107)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)			
Mid-Atlantic	Revenue tax	5,795	-	(368)	-	(1,217)	(26)	1,059	4	1,112	(1)	189	(1)	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189	189		
	Revenue tax	-	-	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	(170)	
Southern	Revenue tax	1,264	-	(15)	(15)	(15)	(15)	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264	1,264		
	Revenue tax	-	-	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	(15)	
Midwest	Revenue tax	1,223	-	(354)	(354)	(354)	(354)	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	
	Revenue tax	-	-	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)
Southwest	Revenue tax	6,778	-	(564)	(564)	(564)	(564)	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	6,778	
	Revenue tax	-	-	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)	(564)
West	Revenue tax	1,223	-	(354)	(354)	(354)	(354)	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223	1,223
	Revenue tax	-	-	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)	(354)

Note: The numbers in parentheses indicate the number of audits that should have been done pursuant to the optimal plan but could not be done because of insufficient staff. The numbers not in parentheses, indicate the extent to which the number of audits that could be done with available staff exceeded the optimal plan.

Those staffing imbalances result in some taxpayers being audited or not audited merely because of the area in which they live. Using the medium nonbusiness class as an example, IRS' revised base plan called for mid-Atlantic region tax auditors to audit the returns of 13,378 more taxpayers than they would have if there were no imbalance and called for western region tax auditors to audit the returns of 40,232 less taxpayers than they would have if there were no imbalance. Put another way, the returns of many taxpayers in the mid-Atlantic region would be audited even though their DIF scores indicated a relatively high degree of compliance while the returns of many taxpayers in the western region would not be audited even though their DIF scores indicated a lesser degree of compliance.

To increase compliance through more audits, IRS often asks the Congress to approve additional audit staff. When the Congress does, IRS tries to correct staffing imbalances.

To account for the additional staff, the national office develops an expansion plan with new figures for the number of returns to be examined for each audit class. Just as for the base plan, DIF scores are used to allocate the expansion plan among the seven regions, with the only difference being that new cutoff scores are computed to account for the additional returns to be audited.

After computing the revised number of audits for each region, the national office then determines in what regions the new employees should be placed to achieve a better balance between the number of audits that should be done in a region and the staff available to do them. In the fiscal year 1975 expansion plan, a better balance was achieved than in the base plan.

Region	Number of direct examination staff years above or below (-) the number that would be needed to meet the optimal plan			
	Revenue agents		Tax auditors	
	Base plan	Expanded plan	Base plan	Expanded plan
North Atlantic	-17.6	-7.6	16.8	1.5
Mid-Atlantic	27.4	-	30.3	.9
Southeast	-18.8	-5.6	-3.1	3.4
Central	-6.0	-4.3	2.9	2.1
Midwest	-	-4.0	15.1	1.1
Southwest	12.2	-2.3	13.7	1.4
Western	-16.2	-6.5	-112.3	a/-47.3

a/ The number of tax auditor direct examination staff-years was still well below optimum in the western region because, according to IRS, the allocation of additional tax auditors to that region was constrained by the ability of the region to recruit, train, and absorb a greater number of tax auditors in one fiscal year.

Before the start of the fiscal year, both the base and expansion regional plans developed by the national office are sent to the regions. These plans are not firm but are proposals which the regions can revise with justification.

#### REGIONAL OFFICE

The mid-Atlantic, southwest, and western regions used differing procedures in developing their fiscal years 1974 and 1975 plans.

#### Mid-Atlantic region

Like the national office, this region generally used DIF scores to allocate both fiscal year plans to the districts. Allocation of low nonbusiness returns to be examined by revenue agents was based on district experience rather than DIF scores because such returns are normally audited by revenue agents only as related pickups. The region did not adjust any of the national office figures before allocating its plans.

Unlike the national office, the region did not adjust its allocations to the district to account for staffing imbalances; it instructed each district to make pro rata adjustments if staff-years available were more or less than called for in the plan.

#### Southwest region

This region also allocated its plans to each district on the basis of DIF scores.

It did not adjust the national office's proposed direct examination staff-years or examination rates before making these allocations, and it did not adjust the allocations for staffing imbalances. The region instructed the districts to make pro rata adjustments for tax auditor imbalances in the 1974 plan, it did not provide any guidance for correcting revenue agent imbalances in the 1974 plan or tax auditor and revenue agent imbalances in the 1975 plan.

#### Western region

This region's procedures differed greatly from those of the mid-Atlantic and southwest regions, and its procedures in fiscal year 1974 differed from those in 1975.

The region did not allocate its fiscal year 1974 plan to the districts. Instead, it advised them of the staff-years available and the examination rates proposed by the national office and asked them to prepare their own plans.

In 1975 the region allocated its plan to the districts. Before doing so, it adjusted the number of direct examination staff-years proposed by the national office apparently because of a different projection of available staff within the region and adjusted the national office's proposed examination rates in all audit classes for revenue agents and in the nonbusiness audit classes for tax auditors. The rates used were the higher of either the regional average or the district office rate experienced during the first 8 months of fiscal year 1974.

The region allocated its plan on the basis of staff distribution within the region. For example, the Los Angeles district had 40.8 percent of the available revenue agent staff-years, and it was allocated 40.8 percent of the revised regional plan for each audit class.

The staff allocation system used by the western region, unlike the DIF allocation system used by the other two regions, does not isolate any district imbalances between staff needed and staff available. We allocated the western region's plan on the basis of DIF scores and found that the Los Angeles district's share under the staff allocation system was about 43,000 returns less than it would have been under the DIF allocation system. Specific examples of regional imbalances follow.

--The Portland district was allocated 2,800 medium nonbusiness returns more and the Los Angeles district was allocated about 13,500 returns less than would have been allocated using DIF scores.

--The Reno and Seattle districts were allocated about 325 and 450 medium business returns more and the Los Angeles district was allocated about 875 returns less than would have been allocated using DIF scores.

Like the national office's plan for the regions, the regions' plans for the districts are not firm but, rather, are proposals to be used by the districts in developing their plans.

## DISTRICT OFFICE

We reviewed the procedures followed by the Baltimore district in the mid-Atlantic region, the New Orleans and Cheyenne districts in the southwest region, and the Los Angeles district in the western region in developing their plans.

Generally, all four districts used data compiled from past audit experience, such as actual examination rates and nonexamination time and considered such other factors as local conditions, available staff, and the experience of the staff.

Los Angeles and New Orleans used local statistics such as dollar yield per return and the no-change rate per audit class to develop their plans. Plans developed using these statistics were, in some cases, inconsistent with national office compliance objectives because staff was shifted from the lesser to the better complying taxpayer classes. For example, in fiscal year 1975 the Los Angeles district reduced the staffing proposed by the region for low business returns, the worst compliance class, by 4.1 direct examination staff-years because it had experienced low yields and high no-change rates in that class. Conversely, the district increased the staffing proposed by the region for medium nonbusiness returns, the best compliance class, by 9.2 direct examination staff-years because that class was one of the best producers in the district.

Cheyenne and New Orleans did not allocate their plans to the examination branches or groups within the district in either 1974 or 1975. Baltimore allocated its 1974 plan to the branch level and its 1975 plan to the branch level for revenue agents and to the group level for tax auditors. Los Angeles allocated its plan to the branch and group levels in both fiscal years.

The national office and two of the three regions allocated their work plans to the next management level on the basis of DIF scores. Neither of the two districts that allocated their plans to the branches or groups used DIF scores to do so; instead, the plans were allocated on the basis of available staff.

In 1974, IRS made a workload study to determine the staffing imbalance at the group (post of duty) level. This study involved comparing the staff available with the staff

needed to examine the number of returns that should be examined. According to mid-Atlantic regional and Los Angeles district officials, however, the results were invalid because of incomplete or erroneous data.

### CONCLUSIONS

Questions often arise as to why a greater percentage of taxpayers in one class or geographical area are audited than in another class or area. The primary reason is that IRS studies and statistics show compliance is worse in some classes or geographical areas than in others.

Although its basic goal is to direct its audit effort where most needed for compliance, IRS is faced with imbalances between the number of audits that should be done and the staff available to do them and must adjust its plan accordingly. Such imbalances result in some taxpayers being audited or not audited merely because of where they reside. The national office and some regions were trying to isolate these imbalances and correct them, but one region included in our review was not. Furthermore, none of the four districts were using DIF scores to isolate imbalances in their posts of duty, although IRS did take an unsuccessful step in that direction in 1974.

DIF scores are available by district and post of duty and should be used to identify compliance problems. The regions and districts would then be in a position to realine their staffs to correct imbalances.

Some districts use no-change and direct yield rates from regular audits in developing their work plans. Such statistics should not be used because (1) returns audited by the districts under their regular audit programs, unlike returns audited under TCMP, are not randomly selected and thus do not reflect compliance for all district taxpayers and (2) plans developed on the basis of such statistics are not in consonance with the national plan.

### RECOMMENDATION TO THE COMMISSIONER OF INTERNAL REVENUE

We recommend that IRS take action to achieve uniformity among regions and districts in developing their work plans. Such uniformity should include the use of DIF scores to identify imbalances between work needed and staff available and should eliminate the use of data which is not representative of compliance characteristics or patterns.

## IRS COMMENTS AND OUR EVALUATION

IRS said that for fiscal year 1977 it would begin phasing in an allocation procedure to provide more uniformity in developing work plans. The allocation model, according to IRS, will include DIF cutoff scores and other criteria to identify workload and staffing requirements for each district on a uniform basis. IRS added that:

--"Fiscal year 1977 represents the first step to adjust the present staffing imbalances but due to budgetary constraints, optimum allocations will not be possible until fiscal year 1978 at the very earliest."

--"The fiscal year 1978 Examination Plan will be developed by the National Office showing the optimum district and regional staffing based on the allocation model. District staffing can then be adjusted and balanced with the workload to the highest degree possible. There may be instances where imbalances could continue beyond fiscal year 1978 where some districts' attrition is so low as to require a longer term phase in to achieve optimum staffing. The only alternative to this would be mandatory transfer of personnel which is an expensive alternative that may have to be faced as the situation occurs."

--"It should be recognized that balanced staffing is not under complete control of the Service. \* \* \* Budget reductions have [an] impact where attrition losses rather than workload control or constrain reallocation capability."

IRS also agreed that no-change and direct-yield data from regular audits should not be used exclusively in developing work plans. It did note, however, that an indication of the effectiveness of resource allocation to the districts could be obtained from their average tax change experience but only in conjunction with other variables, such as staffing limitations and staff-hours per return. Then, according to IRS, allocation of resources to meet workload requirements, as determined by DIF scores, could be supplemented in districts exhibiting marked deviations of yield. We recognize that instances may arise when it would be appropriate to supplement or revise a district's resource allocation on the basis of its yield experience. We would not object to such action provided it was consistent with IRS' long- and short-range goals.

## IMPLEMENTING THE ANNUAL PLAN

IRS establishes long-range goals and develops a plan to meet these goals through a lengthy, systematic process involving many organizational levels. In fiscal year 1974, however, IRS deviated from its goals by using the additional examiners authorized by the Congress to audit returns of taxpayers least needing attention. In effect, IRS concentrated on completing a specific number of audits instead of concentrating on returns from taxpayers with the worst compliance record.

### IRS' commitment to the Congress

On September 17, 1973, the national office instructed each region to expand the fiscal year 1974 plan because the Congress was about to authorize additional revenue agents and tax auditors. These instructions noted that the Commissioner of Internal Revenue "has given us an objective to not only examine more returns than we accomplished in fiscal year 1973, but also to increase the number of examinations in relation to filings, and to do so on a quality basis." IRS justified its budget request for additional staff in 1974 by noting that:

- Many more tax returns would be examined among all classes of taxpayers.
- The program would concentrate on raising the rate of voluntary reporting in low compliance classes while keeping adequate presence in the higher compliance classes and in high yield tax change classes.
- About 306,000 additional tax returns would be examined.
- The additional staff would result in additional tax assessments of about \$250 million.

The Assistant Commissioner for Compliance testified before a Subcommittee of the House Committee on Appropriations that 1/:

---

1/Hearings Before the Subcommittee on Treasury, Postal Service, and General Government of the House Committee on Appropriations, 93d Congress, 1st sess., 522 (1973).

"\* \* \* the projection for fiscal year 1973 is a coverage of 1.9 [percent] for income, estate, and gift tax returns. We intend to examine 1,668,000 returns of all types. With the budget request for fiscal year 1974, we expect to get a coverage of 2.3 [percent], and we will examine close to 2 million returns. This will indicate an increase in both [revenue agent and tax auditor] examinations.

"In connection with [tax auditor] examinations, we will be examining more low business returns."

#### The "big push"

In an effort to meet its commitment to the Congress, the national office issued a memorandum on November 30, 1973, to all regions on ways to meet the expanded 1974 plan. The memorandum listed contingency plans in order of priority that could be "put in motion if you start to fall irreversibly short of planned objectives." The contingency plans involved:

1. Adjusting the interview/correspondence mix for non-business returns examined by tax auditors from the 75/25 objective (but not to exceed 70/30).
2. Reducing the planned tax auditor staff-year applications to the high nonbusiness and the medium and high business classes for the second half of the fiscal year and reallocating them to the remaining nonbusiness classes.
3. Postponing unit III training for tax auditors recruited for fiscal year 1974 until 1975, unless such training was necessary to accomplish the business returns program.
4. Reducing the planned tax auditor staff-year allocations to the individual low business class but only after the adjustment permitted in 2 above had been made to the maximum extent possible. Because of the high noncompliance problem in this class, IRS hesitated to permit this deviation and asked that such action be taken only as a last resort.

This memorandum, in steps two and four, allowed the regions to deviate from the developed work plan for individual returns by directing tax auditor examinations primarily toward the less time-consuming low and medium nonbusiness returns.

IRS accomplishments in fiscal year 1974

IRS met its expanded fiscal year 1974 plan for individual returns by the following percentages.

Percent of plan accomplished (note a)

<u>Audit class</u>	<u>Revenue agents</u>	<u>Tax auditors</u>	<u>Revenue agents and tax auditors combined</u>	<u>1969 voluntary compliance rate</u>
				(percent)
Nonbusiness:				
Low, standard	5,156	117	119	95.2
Low, itemized	187	116	119	88.5
Medium	187	113	120	96.1
High	99	53	86	94.1
Business:				
Low	105	87	92	68.7
Medium	125	125	125	87.8
High	101	44	94	91.2

a/See app. VI for a comparison of the number of audits planned with the number actually done by region in fiscal year 1974.

Our analysis of statistical data showed that:

--Audits of individual returns in fiscal year 1974 increased by 278,100 over fiscal year 1973--232,100 were made by tax auditors (a 20-percent increase over 1973) and 46,000 were made by revenue agents (a 17-percent increase over 1973).

--Of the 232,100 increase in the number of audits by tax auditors, 208,000, or 90 percent, were in the medium nonbusiness class. Interestingly, IRS

statistics show that this class of return has the highest level of voluntary compliance among all the classes of individual returns. Conversely, in the low business class which, according to IRS statistics, has the worst voluntary compliance level by far, the number of audits actually decreased between 1973 and 1974.

According to IRS statistics for fiscal year 1974, it took 2 to 3 times longer to audit a low business return than it did to audit a medium nonbusiness return and audits of medium nonbusiness returns yielded between \$21 and \$46 more per staff-hour than did audits of low business returns. It appears that IRS, to meet its commitments to do more audits and generate additional assessments, directed its audit effort to the areas producing the best yield in the time available.

#### Conclusions

To meet its commitment to the Congress for additional audits and assessments, IRS deviated from its annual plan. Instead of directing its increased audit effort where it would do the most good from a compliance standpoint, IRS directed it to the medium income nonbusiness taxpayer who has historically been the best complier. Such action is counterproductive to achieving IRS' compliance goals.

#### Recommendation to the Commissioner of Internal Revenue

We recommend that IRS avoid further "pushes" to complete audits by refraining from committing itself to a specific number of audits or dollar amount of assessments to justify its request for additional staff.

#### IRS comments

IRS agreed that workload data, in terms of audits or dollars assessed, shown in the budget submission should be considered estimates of objectives that may be accomplished and not objectives that will be accomplished. IRS noted that the fiscal year 1974 congressional budget request used the word "will" in relation to the number of examinations resulting from the increased manpower requested but that the narrative in later years' requests is phrased in terms of "we expect to." IRS noted also that the Commissioner has expressed his desire to consider budget workload data in terms of estimates and not specific commitments.

### THE ANNUAL PLAN AS A QUOTA

Because the annual plan calls for a specific number of audits, it has often been equated with a quota system. IRS has consistently denied that it has a quota for its examiners; others have consistently argued that it does.

We saw no evidence of quotas for individual examiners in the four districts we visited. The work plan was not being allocated to individual examiners, and there were no indications that examiners were being told to do a specific number of audits. Although we saw no evidence of quotas, the question remains whether examiners feel pressured to complete as many audits as possible and whether that pressure, if any, is unreasonable.

We asked examiners and group managers what factors had the greatest bearing on performance evaluations. The factors examiners cited most were case time or number of cases closed, taxpayer relations, and audit quality. Group managers never referred to number of cases closed as a factor but often referred to case time. Although audit quality is considered important, some examiners felt that the greatest emphasis is on completing cases and a few said this sometimes results in potentially lucrative areas going unaudited and sometimes precludes more audit emphasis on unreported income.

Because IRS' voluntary compliance goals are predicated on accomplishing a certain number of audits, it must be concerned that examiners spend enough time to cover the most significant audit issues (hereinafter referred to as a quality audit) while not wasting time on insignificant matters. While IRS cannot be faulted for expecting its examiners to adhere to reasonable time constraints, some examiners apparently believe that they are being pressured to adhere to unreasonable constraints and that IRS is concerned with quantity to the detriment of quality. Certain facts, like the push to complete audits in 1974, lend credence to such belief. In addressing the issues of quantity versus quality and unreasonable time constraints, we do not believe that an annual plan, in and of itself, is the problem. A realistic plan can provide for a specific number of audits without sacrificing quality. The more appropriate question then is not whether IRS should have an annual plan but whether the plan it has is realistic.

The basic factors used in developing the annual plan are the examination rates--the number of examinations that can be done in a particular audit class during a direct examination staff-year. If these rates, which are based on past experience, realistically reflect the average time it takes to make a quality audit, then there should be no conflict between quantity and quality and the examiners should not feel pressured.

Problems arise when IRS deviates from its plan as it did in 1974 and as it intended to do again in 1975 as indicated by the following excerpt from a national office memorandum to each regional office.

"While your plan has been approved, the plans for district offices, on a National basis, are about 40,000 returns below the Commissioner's stated objectives for FY 1975 or two percent of plan. We have no reservations that the additional number of examinations can be accomplished within the present plan concept. Instead of requesting a further revision of the plan, we will monitor on the basis of the Commissioner's objective anticipating 102 percent accomplishment of total district plans."

To accomplish 102 percent of plan, IRS would have to make more examinations per direct examination staff-year, which would distort the examination rates that will be used in developing future annual plans. This distortion adds to the distortion caused by the deviation from the 1974 plan, and any subsequent deviations will further compound the distortion. Thus, examination rates that may have been reasonable when first developed could very likely have become unreasonable. The examiner then faces a dilemma--the desire to do a quality audit versus management's desire to meet a plan that is based on examination rates that do not allow sufficient time for a quality audit.

To resolve this dilemma and be in a better position to justify its annual plan and the resultant time constraints, IRS should evaluate the reasonableness of its examination rates and avoid actions that tend to distort those rates. This would seem to require a controlled study, for each audit class, of what constitutes a quality audit and how long, on the average, it takes to do one.

The results of such a study could also be used as a basis for conveying to examiners (1) what management considers to be a quality audit; that is, the depth of audit that management expects them to achieve, and (2) how the examination rates developed by management tie into those expectations. The results, in effect, could be used as a basis for allaying examiners' concerns about quantity versus quality.

Because examination rates are based on the average time to do an audit, deviations from that rate caused by audit complexity and examiner experience are to be expected. A particular average may be unachievable by certain examiners or groups of examiners while easily surpassable by others. Thus, by allocating the plan down to the group level, as is done in some districts, IRS could be placing unreasonable pressure on some groups.

We recognize the need for allocating the annual work plan to districts and even in some cases to branches, especially those branches that have a diversified examination workload. We fail to see, however, a compelling need to allocate the plan to the group level, as some districts do. The only apparent reason for doing so is to assist the groups in monitoring their progress. In our opinion, the returns program manager could take care of such monitoring on a districtwide basis as provided for in the IRS manual.

#### Recommendations to the Commissioner of Internal Revenue

We recommend that IRS conduct a controlled study to evaluate the reasonableness of its examination rates and that it discontinue the practice in some districts of allocating the annual plan to groups.

#### IRS comments and our evaluation

IRS agreed that annual audit plans should not be allocated to the group level; according to IRS, elimination of this practice was discussed at a June 1976 meeting of IRS' Assistant Regional Commissioners responsible for audit.

IRS did not agree with the need for a controlled study to evaluate the reasonableness of examination rates; it had an alternative approach. As we understand from discussions with Audit Division personnel, IRS' approach, used in developing its fiscal year 1977 revenue agent examination

rates, basically involved using the past 4 years' experience, eliminating the aberrational year, and computing an average for the remaining 3 years for each audit class in each district. The average district rates for each audit class were then adjusted again to further eliminate aberrations--the result being a standard rate for each audit class. A productivity improvement factor was then added to any district average rate that fell below the standard; no improvement factor was added to an average rate that was above the standard. Under this approach, according to IRS, districts with higher average rates will be permitted to concentrate on improved audit quality while districts with lower rates will have to improve their efficiency.

This approach should produce more reasonable revenue agent examination rates and, in conjunction with eliminating the allocation of the annual plan down to the groups, may help alleviate the apparently unreasonable pressure being felt by some examiners to complete cases. We are not convinced, however, that the newly computed rates accurately reflect the time needed to conduct a quality audit since there is no assurance that examiners, in any of the past 4 years, were ever afforded adequate time to do such an audit. In other words, the newly computed rates might just be averages of previously unreasonable rates. Also, IRS' new approach for computing examination rates only applies to revenue agents. According to IRS, tax auditor examination rates for 1977 were computed by adding an improvement factor to the average rate achieved by all tax auditors in 1976.

We still believe that a controlled study is the only way to insure reasonable examination rates. Until such a study is undertaken, we believe that IRS, as a minimum, should (1) convey to its examiners what management considers to be a quality audit; that is, the depth of audit that management expects them to achieve and how its examination rates tie into those expectations and (2) solicit and recognize the comments of examiners on the reasonableness of the examination rates.

Department of the Treasury / Internal Revenue Service / Washington, D.C. 20224

## Commissioner

AUG 10 1976

Mr. Victor Lowe  
Director, General Government Division  
U. S. General Accounting Office  
Washington, D. C. 20224

Dear Mr. Lowe:

We appreciate the opportunity to review your draft report to the Joint Committee on Internal Revenue Taxation entitled, "Selection of Individual Income Tax Returns for Audit by the Internal Revenue Service."

The report discloses that tests conducted by IRS and GAO show that the selection system, including manual screening, is generally effective. However, the report indicates a number of areas where improvements in the system can be made. In most instances, we are in agreement with these recommendations. Our comments regarding specific recommendations are enclosed with explanations in those situations where some disagreement exists. These are referenced to the applicable page number in the digest and report. Also, we noted a number of editorial changes which we feel should be made in the final report. The changes are listed in Attachment A. [GAO notes 1 and 2]

We have also enclosed a current list of unallowable items (Attachment B) for your information and have marked with an asterisk, those new items that do not appear in Appendix I of the report. In addition, two items (Job Seeking Expenses and Self-Employment Tax) are no longer included in the unallowable items program. However, at the time GAO began its audit, the list of [GAO note 3] unallowable items as contained in Appendix I was complete and accurate.

Finally, we would like to explain the voluntary compliance level (VCL) computations included in the report. The VCL, estimated by TCMP, used by the Service at the time of the GAO report was computed by dividing total tax reported on returns filed by the sum of total tax reported on returns filed plus tax change. In computing "tax change", total deficiencies (i. e., underpayments of tax) were added, not netted, to total overassessments (i. e., overpayments of tax). Recently, we reassessed this computation because it treated taxpayers who overpaid their tax as "noncompliant". We do not feel these taxpayers should be considered noncompliant for the purposes for which the VCL figure was used. Therefore, we have revised our computation so that taxpayers in a filing population who overpay their tax are treated as "compliant" for these purposes. Of course, we still recognize that

these taxpayers have made an error on their return, and that we have a responsibility to make refunds and to help them avoid errors in the future. Attachment C includes our revised voluntary compliance levels estimated from TCMP data and an explanation of the new equation being used. In this attachment, it appears that the estimated voluntary compliance levels for all returns have been declining from one TCMP survey to the next. However, when sampling variability is taken into consideration, these declines (differences) are not significant at the 95% probability level (see Attachment D). Only the differences for two audit classes are significant depending on the years compared. [GAO note 4]

We would appreciate the opportunity of meeting with you to discuss our comments on this report. My assistant, Tom Glynn, will follow up on this to arrange a meeting.

With kind regards,

Sincerely,



Commissioner

Enclosures

- GAO notes:
1. Some of IRS' comments have been deleted because they pertained to points discussed in the draft report but dropped from the final report or because they were not directly pertinent to our findings, conclusions, or recommendations.
  2. IRS' editorial comments are not included as part of this appendix. The suggested changes have been incorporated in the report, where appropriate.
  3. The current list of unallowable items is included in appendix II.
  4. Attachments C and D to the Commissioner's letter are not included in this appendix. The report has been revised to reflect IRS' change in computing voluntary compliance levels.

Page references in IRS' comments may not correspond to pages in the final report.