



*Improvements in Return Scoring and  
Resource Allocation at the Strategic Level  
Could Enhance Examination Productivity*

**April 24, 2019**

**Reference Number: 2019-30-024**

This report has cleared the Treasury Inspector General for Tax Administration disclosure review process and information determined to be restricted from public release has been redacted from this document.

**Redaction Legend:**

1 = Tax Return/Return Information

---

Phone Number / 202-622-6500

E-mail Address / [TIGTACommunications@tigta.treas.gov](mailto:TIGTACommunications@tigta.treas.gov)

Website / <http://www.treasury.gov/tigta>



**To report fraud, waste, or abuse, call our toll-free hotline at:**

**1-800-366-4484**

**By Web:**

**[www.treasury.gov/tigta/](http://www.treasury.gov/tigta/)**

**Or Write:**

Treasury Inspector General for Tax Administration  
P.O. Box 589  
Ben Franklin Station  
Washington, D.C. 20044-0589

Information you provide is confidential and you may remain anonymous.



## HIGHLIGHTS

### IMPROVEMENTS IN RETURN SCORING AND RESOURCE ALLOCATION AT THE STRATEGIC LEVEL COULD ENHANCE EXAMINATION PRODUCTIVITY

## Highlights

Final Report issued on April 24, 2019

Highlights of Reference Number: 2019-30-024 to the Commissioner of Internal Revenue.

### IMPACT ON TAXPAYERS

The IRS's primary objective in selecting returns for examination is to promote the highest degree of voluntary compliance. This requires the exercise of professional judgement in selecting returns to assure all taxpayers of equitable consideration and in making the most efficient use of staffing resources. The IRS uses a variety of sources to select returns for audit and strives to select those returns for which its examiners are likely to find areas of noncompliance.

### WHY TIGTA DID THE AUDIT

This audit was initiated to evaluate the Small Business/Self-Employed Division's strategic priority selection methods used to identify individual income tax returns for examination by revenue agents and to assess the effectiveness of the strategic priorities with an emphasis on the Discriminant Function (DIF) selection method.

### WHAT TIGTA FOUND

Between Fiscal Years 2014 and 2016, the Small Business/Self-Employment Division Examination function (Examination function) used 10 strategies to group similar examination work. During each of these fiscal years, the Examination function established planned examined returns closed (hereafter referred to as closures) at the strategic priority level. However, the Examination function does not establish performance goals at the strategic level; rather, it uses cumulative direct examination staff years and cumulative closures combined for revenue agents and tax

compliance officers to determine whether the examination plan was met.

TIGTA found the Examination function did not take corrective action when actual closures exceeded or fell behind planned closures for a specific strategic priority. As a result, the IRS lost the opportunity to assess an additional \$262.5 million on individual income tax return examinations by revenue agents for Fiscal Year 2016.

TIGTA also reviewed the DIF scores for individual returns filed with the IRS during Calendar Year 2015, and examined and closed by revenue agents and tax compliance officers through December 2017. Even though the DIF identifies returns with examination potential, TIGTA found that, more often than not, the highest DIF score individual returns examined by revenue agents do not necessarily result in a higher net tax assessment for most examination classes.

### WHAT TIGTA RECOMMENDED

TIGTA recommended that 1) the IRS establish actionable performance goals, and monitor and take corrective action specific to staffing resources and closures for each strategic priority to ensure that it provides balanced examination coverage, and 2) determine why the highest DIF scores for individual returns examined by revenue agents and tax compliance officers did not result in higher net tax assessments than lower DIF scores for most examination classes.

The IRS disagreed with the recommendations. The management response stated that Examination does not control how its examination plan is worked. It also stated that that National Research Program data supports its DIF models. It does not believe that the specific examination results in the report provide any actionable information relevant to the current DIF models. TIGTA believes that the IRS does control how the plan is worked and should make adjustments as needed to meet the plan. Moreover, the examination results should be used to improve the DIF.



TREASURY INSPECTOR GENERAL  
FOR TAX ADMINISTRATION

DEPARTMENT OF THE TREASURY  
WASHINGTON, D.C. 20220

April 24, 2019

**MEMORANDUM FOR COMMISSIONER OF INTERNAL REVENUE**

**FROM:** Michael E. McKenney  
Deputy Inspector General for Audit

**SUBJECT:** Final Audit Report – Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity (Audit #201730031)

This report presents the results of our review to perform trend analyses of the Small Business/Self-Employed Division's strategic priority selection methods used to identify individual income tax returns for examination by revenue agents and assess the effectiveness of the strategic priorities with an emphasis on the Discriminant Function selection method. This audit was included in our Fiscal Year 2018 Annual Audit plan and addresses the major management challenge of Improving Tax Compliance.

Management's complete response to the draft report is included as Appendix VII.

Copies of this report are also being sent to the Internal Revenue Service managers affected by the report recommendations. If you have any questions, please contact me or Matthew A. Weir, Assistant Inspector General for Audit (Compliance and Enforcement Operations).



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

## *Table of Contents*

<a href="#">Background</a> .....	Page 1
<a href="#">Results of Review</a> .....	Page 7
<a href="#">Conducting More Examinations Than Planned in Some Strategies and Fewer in Others Negatively Impacts Compliance Results</a> .....	Page 7
<a href="#">Recommendation 1:</a> .....	Page 13
<a href="#">Improved Processes Have Led to Reductions in Surveyed and No-Change Examination Closures by Revenue Agents of Individual Returns for Most Strategies</a> .....	Page 14
<a href="#">There Are Opportunities to Improve Discriminant Function Scoring of Individual Returns</a> .....	Page 21
<a href="#">Recommendation 2:</a> .....	Page 31
<b>Appendices</b>	
<a href="#">Appendix I – Detailed Objective, Scope, and Methodology</a> .....	Page 32
<a href="#">Appendix II – Major Contributors to This Report</a> .....	Page 35
<a href="#">Appendix III – Report Distribution List</a> .....	Page 36
<a href="#">Appendix IV – Outcome Measure</a> .....	Page 37
<a href="#">Appendix V – Results of Individual Primary Return Examinations by Revenue Agents and Tax Compliance Officers</a> .....	Page 39
<a href="#">Appendix VI – Glossary of Terms</a> .....	Page 43
<a href="#">Appendix VII – Management’s Response to the Draft Report</a> .....	Page 45



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

## *Abbreviations*

AIMS	Audit Information Management System
ATTI	Abusive Transactions and Technical Issues
DAO	Discriminant Function and All Other
DESY	Direct Examination Staff Year
DIF	Discriminant Function
FY	Fiscal Year
GAO	Government Accountability Office
HIHW	High Income and High Wealth
IRM	Internal Revenue Manual
IRS	Internal Revenue Service
No.	Number
NRP	National Research Program
OFF	Offshore
OPP	Other Priority Programs
P&RP	Promoters and Return Preparer Strategies
PP&A	Performance, Planning, and Analysis
PSP	Planning and Special Programs
RAAS	Research, Applied Analytics, and Statistics Compliance Modeling Lab
SB/SE	Small Business/Self-Employed
SEP&F	Special Enforcement and Fraud Programs
TIGTA	Treasury Inspector General for Tax Administration
U&UI	Unreported and Underreported Income



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

### *Background*

The average annual gross Tax Gap for Tax Years 2008 through 2010 is estimated to be \$458 billion.<sup>1</sup> The largest component, \$387 billion, is attributable to underreporting of taxes (most of this, \$264 billion, is associated with underreporting of individual income taxes). The net Tax Gap is estimated to be \$406 billion.<sup>2</sup> To address underreporting by taxpayers, the Internal Revenue Service (IRS) selects tax returns for examination. The IRS's Small Business/Self-Employed (SB/SE) Division examines individual tax returns and business tax returns with assets under \$10 million. The primary objective in selecting returns for examination is to promote the highest degree of voluntary compliance on the part of taxpayers. This requires the exercise of professional judgment in selecting returns to assure all taxpayers of equitable consideration and in making the most efficient use of staffing resources. The SB/SE Division, Examination, Performance, Planning, and Analysis (PP&A) function prepares the annual SB/SE Division examination plan. In implementing the plan, the SB/SE Division's strategy is to have coverage across return categories that fall within the purview of the SB/SE Division.<sup>3</sup> The IRS uses a variety of sources to select returns for audit, and strives to select for audit those returns for which its examiners are likely to find areas of noncompliance and recommend changes to one or more items reported on the return.<sup>4</sup>

Between Fiscal Years (FY) 2014 and 2016, the SB/SE Division Examination function used 10 strategic priorities to group similar examination work.<sup>5</sup> SB/SE Division Examination function management explained that the concept behind the strategies is to categorize and distribute work by type in order to maintain a balance. According to the PP&A function, the Strategic Priority Coding List is a tool used to avoid duplicate counting of examined returns closed (hereafter referred to as closures) and examination results. The list identifies various codes including Source Codes, Project Codes, and Examination Return Control System Tracking Codes that are unique to a strategy. Each strategy provides a different type of inventory. IRS management defined these strategies as follows:

---

<sup>1</sup> See Appendix VI for a glossary of terms.

<sup>2</sup> IRS Publication 1415, *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008–2010*, pp. 1–2 (May 2016).

<sup>3</sup> Internal Revenue Manual 4.1.1.2.3(1) (Oct. 25, 2017).

<sup>4</sup> Unless otherwise stated, the scope of this review was limited to SB/SE Division examinations of individual income tax returns conducted by revenue agents in field offices in the seven SB/SE Division examination areas. References to the “examination plan” throughout this report are to the SB/SE Division Field Examination Plan. The SB/SE Division has seven Field Examination Areas based on geographical location. They are North Atlantic, South Atlantic, Central, Midwest, Gulf States, Western, and Southwest.

<sup>5</sup> Hereafter referred to as “strategy” when used singularly or as “strategies” when used in the plural.



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

- The National Research Program (NRP) strategy provides a statistically valid random sample of filed returns representative of the compliance characteristics of taxpayers. Returns in this strategy are assigned to examiners as quickly as possible, and surveys before or after assignment are limited.<sup>6</sup>
- The Promoters and Return Preparer Strategies (P&RP) include examining returns of clients of return preparers, examining returns of clients of promoters who also prepare returns, examining promoter returns, examining return preparer tax returns, assessment of preparer penalties, and pursuing injunctions against abusive preparers and promoters.
- The Offshore (OFF) strategy addresses returns that appear to have abusive issues related to offshore accounts and/or international issues.
- The Abusive Transactions and Technical Issues (ATTI) strategy focuses on domestic abusive transactions and tactics.
- The High Income and High Wealth (HIHW) strategy addresses HIHW taxpayers including individuals who did not file a tax return.<sup>7</sup>
- The Unreported and Underreported Income (U&UI) strategy involves indications of unreported income.
- The Special Enforcement and Fraud Programs (SEP&F) strategy addresses unreported income that may be criminal in nature.
- The Other Priority Programs (OPP) strategy contains specific work streams and important projects that are specifically coded and do not fit into the other strategies.
- The Discriminant Function (DIF) strategy contains returns selected based on algorithms. The results of NRP examinations are used to update the algorithms.
- The All Other strategy includes all other work that is not included in the other strategies, but does not contain specific work streams.

The annual SB/SE Division's examination plan details resource allocation of revenue agents and tax compliance officers for each strategic priority. Figure 1 shows the SB/SE Division Examination U.S. Program Monitoring report for revenue agent individual returns planned,

---

<sup>6</sup> Internal Revenue Manual 4.22.4.2.3(1) and 4.22.4.3(1) (June 6, 2018).

<sup>7</sup> For FY 2018, projects under the HIHW High Income Nonfiler subcategory were moved to Strategy Number 6 (Nonfiler), which was previously a placeholder, and projects under the HIHW subcategory were moved to Strategy Number 9 OPP with other national Compliance Initiative Projects. The former HIHW strategy became a placeholder on the FY 2018 Priority Coding List containing no projects.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

actual closures, and the percentage of the plan completed through September for FYs 2014 through 2016.<sup>8</sup>

**Figure 1: Revenue Agent Individual Returns  
Planned Versus Actual Closures by Strategy for FYs 2014 Through 2016**

Strategy	FY 2014			FY 2015			FY 2016		
	Planned	Actual	% of Plan <sup>9</sup>	Planned	Actual	% of Plan	Planned	Actual	% of Plan
NRP	9,756	10,135	104%	10,600	10,418	98%	13,061	16,904	129%
P&RP	5,473	5,270	96%	4,204	3,511	84%	2,609	2,601	100%
OFF	19,983	16,982	85%	18,274	17,390	95%	16,149	17,059	106%
ATTI	5,783	3,181	55%	2,464	2,957	120%	3,104	2,227	72%
HIHW	6,203	5,757	93%	4,612	5,274	114%	4,855	4,908	101%
U&UI	1,902	4,699	247%	5,012	8,593	171%	6,249	9,802	157%
SEP&F	2,498	1,979	79%	1,893	1,874	99%	2,102	1,682	80%
OPP	12,651	15,751	125%	13,289	14,955	113%	14,693	14,651	100%
DIF	N/A	46,088	N/A	N/A	39,439	N/A	N/A	25,834	N/A
All Other	N/A	11,828	N/A	N/A	8,514	N/A	N/A	6,675	N/A
DAO <sup>10</sup>	54,934	57,916	105%	54,737	47,953	88%	40,501	32,509	80%
<b>Total Closures</b>	<b>119,183</b>	<b>121,670</b>	<b>102%</b>	<b>115,085</b>	<b>112,295</b>	<b>98%</b>	<b>103,323</b>	<b>102,343</b>	<b>99%</b>

Source: Treasury Inspector General for Tax Administration (TIGTA) summary of SB/SE Division Examination U.S. Program Monitoring Reports FY 2014 through FY 2016.

Figure 1 shows planned revenue agent individual return closures for eight of the 10 strategies and combined planned closures for the DIF and All Other (DAO) strategies. The IRS did not report these planned closures separately for the DAO strategies in FYs 2014 to 2016. IRS management stated that the DIF strategy is built into the examination plan after allocations have

<sup>8</sup> We relied on the data provided by the IRS, including the U.S. Program Monitoring, U.S. Monthly Performance, SB/SE Division’s Director Field Exam Briefings, or the U.S. Strategy Closures and New Starts reports, and did not independently validate the IRS’s information. Fiscal year cumulative data in the U.S. Program Monitoring Reports are as of the month end and may differ from final Audit Information Management System data. Figures in the U.S. Program Monitoring Reports are as of September of the applicable fiscal year.

<sup>9</sup> The Percentage of Plan column is a calculated figure determined by dividing the number of actual closures by the number of planned closures.

<sup>10</sup> The DAO strategy is the combined figure for the DIF and All Other strategies. The source of the data for this figure, the U.S. Program Monitoring report, does not report individually the planned closures for the DIF and All Other strategies.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

been made to the other strategies. According to SB/SE Division management, the purpose of the strategies is not based solely on whether they are productive in terms of dollars assessed. For example, the compliance data provided by the examinations of returns selected for the NRP strategy update the DIF formulas, assist the IRS in estimating the Tax Gap, and assist the IRS in meeting its obligations under the Improper Payments Information Act of 2002 and subsequent legislation and guidance for annual reporting on improper payments of refundable tax credits.<sup>11</sup> The P&RP strategy can be productive because it results in additional assessments; however, it is primarily designed to correct the actions of tax preparers who are not following proper tax preparation practices.

The development of the SB/SE Division's examination plan involves allocating available time to the strategies. The time allocated to the DIF strategy is the residual time after applying time to the other strategies and compliance activities. Figure 2 shows that the Direct Examination Staff Years (DESY) expended by revenue agents on examinations of individual income tax returns has declined from 1,827.33 in FY 2014 to 1,608.02 in FY 2016. During that same time, the number of the DESYs expended on the NRP strategy more than doubled from 149.47 to 353.97 DESYs.

---

<sup>11</sup> Pub. L. No. 107-300, 116 Stat. 2350.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 2: Planned and Actual Direct Examination Staff Years for Revenue Agent Examinations of Individual Returns by Strategy for FYs 2014 Through 2016<sup>12</sup>**

Strategy	FY 2014			FY 2015			FY 2016		
	Planned DESYs	Actual DESYs	% of Plan	Planned DESYs	Actual DESYs	% of Plan	Planned DESYs	Actual DESYs	% of Plan
NRP	218.93	149.47	68%	340.09	385.81	113%	323.20	353.97	110%
P&RP	68.32	57.43	84%	57.22	45.30	79%	34.64	37.57	108%
OFF	128.89	103.57	80%	118.69	98.11	83%	123.91	107.53	87%
ATTI	88.80	37.54	42%	38.65	37.42	97%	47.12	26.43	56%
HIHW	79.47	89.74	113%	60.61	74.57	123%	54.71	66.18	121%
U&UI	41.96	110.30	263%	106.64	155.69	146%	101.13	164.62	163%
SEP&F	47.79	32.15	67%	34.36	28.73	84%	37.77	25.70	68%
OPP	210.48	259.77	123%	217.08	249.94	115%	264.27	264.22	100%
DIF	N/A	829.85	N/A	N/A	607.50	N/A	N/A	470.38	N/A
All Other	N/A	157.51	N/A	N/A	114.30	N/A	N/A	91.42	N/A
DAO <sup>13</sup>	989.96	987.36	100%	934.85	721.81	77%	698.89	561.80	80%
<b>Total DESYs</b>	<b>1,874.60</b>	<b>1,827.33</b>	<b>97%</b>	<b>1,908.19</b>	<b>1,797.37</b>	<b>94%</b>	<b>1,685.64</b>	<b>1,608.02</b>	<b>95%</b>

Source: TIGTA analysis of SB/SE Examination Division U.S. Program Monitoring Reports FYs 2014 through 2016.

This review was performed with information obtained from the Research, Applied Analytics, and Statistics, Compliance Modeling Lab (hereafter collectively referred to as the RAAS) located in Washington, D.C., and in the SB/SE Division Examination Case Selection function, the SB/SE Division Performance, Planning and Analysis function, and with SB/SE Examination Division executives in various locations during the period July 2017 through June 2018. This review focused on the selection of examinations of individual taxpayer returns conducted by SB/SE Division revenue agents in the field offices. We relied primarily on interviews with IRS personnel and reviews of available documentation, such as status reports, coding documents, and examination plans. We conducted this performance audit in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objective. We believe that the evidence obtained provides a

<sup>12</sup> Figures in the U.S. Program Monitoring Reports are as of September of the applicable fiscal year.

<sup>13</sup> The DAO strategy is the combined figure for the DIF and All Other strategies. The source of the data for this figure, the U.S. Program Monitoring Report, does not report individually the planned DESYs for the DIF and All Other strategies.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

reasonable basis for our findings and conclusions based on our audit objective. Detailed information on our audit objective, scope, and methodology is presented in Appendix I. Major contributors to the report are listed in Appendix II.



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

*Results of Review*

**Conducting More Examinations Than Planned in Some Strategies and Fewer in Others Negatively Impacts Compliance Results**

By significantly deviating from its planned use of resources at the strategic priority level for examinations of individual returns by revenue agents, the IRS likely significantly reduced the total amount assessed. As previously stated, the SB/SE Division's approach is to provide coverage across return categories that fall within its purview. The examination plan details resource allocation of revenue agents and tax compliance officers for each strategic priority. However, the IRS does not measure its adherence to the planned allocation of resources. The SB/SE Division limits its measurement to two planned targets 1) cumulative DESYs and 2) cumulative closures, both of which are collective in nature by resource types, *i.e.*, revenue agents, tax compliance officers, and combined results.

According to the Government Accountability Office (GAO), performance measurement is the ongoing monitoring and reporting of program accomplishments, particularly progress toward pre-established goals.<sup>14</sup> Performance measurement focuses on whether a program has achieved its objectives, expressed as measurable performance standards.<sup>15</sup> The PP&A function prepares reports for monitoring the examination plan, among them are the U.S. Monthly Performance Report, the U.S. Program Monitoring Report, and the SB/SE Division's Director Field Exam Briefing (hereafter referred to as Briefing). The PP&A function provides these reports to SB/SE Division Examination function management.

The GAO's *Standards for Internal Control in the Federal Government* provides "Monitoring" as one of the five components for internal control. Monitoring provides that management should establish and operate monitoring activities to assess the quality of performance and promptly take corrective actions in order to achieve objectives. The IRS has established monitoring activities at the strategic priority level, but does not take prompt corrective action because these metrics are only for diagnostic purposes. Failure to take prompt corrective action may result in the failure to select sufficient returns of all classes of returns, resulting in inequitable consideration of all classes of returns and inefficient use of resources and staffing.

---

<sup>14</sup> The Merriam-Webster dictionary defines goal as "the end toward which effort is directed." Similarly, target is defined as "a goal to be achieved," while it defines metric as a "standard of measurement." All three terms are used throughout this report.

<sup>15</sup> GAO, GAO-11-646SP, *Performance Measurement and Evaluation: Definitions and Relationships*, p. 2 (May 2011).



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**The U.S. Monthly Performance Report does not include performance by strategy**

As shown in Figure 3, the U.S. Monthly Performance Report summarizes some performance goals for revenue agents, tax compliance officers, and combined results established within the SB/SE Division examination plan. The report shows historical data for two fiscal years and current fiscal year data. However, it does not include performance by strategy.

**Figure 3: U.S. Monthly Performance Report, Overall Performance Results for FYs 2014 Through 2016**

Metric	Historical Data <sup>16</sup>		FY 2016 Data <sup>17</sup>			
	FY 2014 Actual	FY 2015 Actual	FY 2016 Actual	FY 2016/ FY 2015 % Change	FY 2016 Target	% of Target
<b>Closures</b>						
Revenue Agent Closures	171,558	160,984	144,363	(10%)	143,618	100.5%
Tax Compliance Officer Closures	133,990	125,440	116,637	(7%)	110,261	105.8%
<b>Combined Results Closures</b>	<b>305,548</b>	<b>286,424</b>	<b>261,000</b>	<b>(9%)</b>	<b>253,879</b>	<b>102.8%</b>

Source: TIGTA analysis of SB/SE Division Examination U.S. Monthly Performance Report September FY 2016.

Figure 3 shows that closures have decreased from FYs 2014 to 2016 for revenue agent, tax compliance officer, and combined results. The report also computes the percentage of target by comparing the FY 2016 target to date to the FY 2016 actual to date results for revenue agent, tax compliance officer, and combined results. In the FY 2016 Briefings, the total number of closures was emphasized as the metric used to determine whether the SB/SE Division examination plan was met. According to IRS management, the total number of closures is the only metric used to determine whether the SB/SE Division examination plan was met. However, as previously stated, there are two planned targets, *i.e.*, 1) cumulative DESYs and 2) cumulative closures. We asked if there is an established range for determining if they met the established goal or not. IRS management responded that the acceptable range for determining whether the overall plan was met changes from year to year, depending on the challenges faced, with 95 percent through 105 percent being the normal range. Data regarding the strategies are not included on the U.S. Monthly Performance Report. Furthermore, the SB/SE Division Examination function does not

<sup>16</sup> Historical data in the U.S. Monthly Performance Reports are based on Audit Information Management System data.

<sup>17</sup> Fiscal year cumulative data in the U.S. Monthly Performance Reports are as of the month end and may differ from Audit Information Management System data.



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

use strategic priority-level accomplishments when measuring the success of the examination plan.

**The U.S. Program Monitoring Report includes strategy-level information and shows some strategies exceeded or fell behind goals for individual examination closures by revenue agents**

Like the U.S. Monthly Performance Report, the U.S. Program Monitoring Report summarizes some performance targets for revenue agents, tax compliance officers, and combined results established within the SB/SE Division examination plan. The report provides detailed information on the performance of the plan by strategy. To determine whether the SB/SE Division has established performance goals to assess the effectiveness of each strategy for individual income tax return examinations by revenue agents, we reviewed the September U.S. Program Monitoring Reports for FYs 2014 through 2016, as reflected in Figure 4.<sup>18</sup> The SB/SE Division summarizes planned performance goals for the DESYs and closures on its U.S. Program Monitoring Report.

However, IRS management stated the metrics at the strategy level are not for measuring program performance. Rather, the metrics are for diagnostic purposes only. IRS management defined “diagnostic purposes only” as reviewing the other metrics to determine the likelihood that SB/SE Division examination plan closures will be met and to identify factors that contribute to the plan or factors for which improvements can be made. The U.S. Program Monitoring Report also shows the actual closures to date and computes a percentage of plan by comparing the actual closures to date to the planned closures to date. Because SB/SE Division management established an acceptable range for the overall SB/SE Division examination plan, we used the same acceptable range for FYs 2014 through 2016 to determine if revenue agent individual income tax return closures by strategy met the planned closures as shown in Figure 4.<sup>19</sup>

---

<sup>18</sup> Fiscal year cumulative data in the U.S. Program Monitoring Reports are as of the month end and may differ from Audit Information Management System data.

<sup>19</sup> The plan is “Met” if actual closures fall between 95 percent and 105 percent of the planned closures, the plan is “Under” if the actual closures fall below 95 percent of the planned closures, and the plan is “Over” if the actual closures are above 105 percent of the planned closures.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 4: Revenue Agent Individual Returns Planned Versus Actual Closures Over/Under Analysis by Strategy for FYs 2014 Through 2016**

Strategy	FY 2014			FY 2015			FY 2016		
	Actual Closures <sup>20</sup>	% of Plan	Plan Met	Actual Closures	% of Plan	Plan Met	Actual Closures	% of Plan	Plan Met
NRP	10,135	104%	Met	10,418	98%	Met	16,904	129%	Over
P&RP	5,270	96%	Met	3,511	84%	Under	2,601	100%	Met
OFF	16,982	85%	Under	17,390	95%	Met	17,059	106%	Over
ATTI	3,181	55%	Under	2,957	120%	Over	2,227	72%	Under
HIHW	5,757	93%	Under	5,274	114%	Over	4,908	101%	Met
U&UI	4,699	247%	Over	8,593	171%	Over	9,802	157%	Over
SEP&F	1,979	79%	Under	1,874	99%	Met	1,682	80%	Under
OPP	15,751	125%	Over	14,955	113%	Over	14,651	100%	Met
DAO	57,916	105%	Met	47,953	88%	Under	32,509	80%	Under
<b>Total Closures<sup>21</sup></b>	<b>121,670</b>	<b>102%</b>	Met	<b>112,925</b>	<b>98%</b>	Met	<b>102,343</b>	<b>99%</b>	Met

Source: TIGTA analysis of SB/SE Division Examination September U.S. Program Monitoring Reports FYs 2014 through 2016.

Based on our analysis, the SB/SE Division met planned closures in three strategies (P&RP, HIHW, and OPP), exceeded planned closures in three strategies (NRP, OFF, and U&UI), but did not meet planned closures in four strategies (ATTI, SEP&F, and DAO combined), although the SB/SE Division met its Total Closures goal for FY 2016. According to the RAAS, two factors related to the Tax Year 2012 NRP sample contributed to the increase in NRP return closures in FY 2016. The sample size was larger than expected, and return sample cases were delivered late.

To determine the impact of not following the planned individual income tax return closures by strategic priority for revenue agents, TIGTA analyzed the impact of the strategies that exceeded or fell below planned metrics using the “dollars per return” metric contained in the FY 2016 September U.S. Program Monitoring Report. TIGTA applied the IRS’s acceptable range used to measure whether its annual targeted closures were met, *i.e.*, where 95 percent through 105 percent would be considered as meeting the plan. We found the following:

<sup>20</sup> Actual Closures and Percentage of Plan data are from Figure 1.

<sup>21</sup> The Total Closures and corresponding percentages in Figure 4 will not match the closures and corresponding percentages in Figure 3 because Figure 4 data are limited to individual returns, while Figure 3 data include individual, business, and employment tax returns.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

- Four strategies (ATTI, SEP&F, and DAO combined) fell below the planned metric, and had it been followed could have potentially resulted in 9,289 additional closures with \$356.8 million in potential additional assessments.
- Three strategies (NRP, OFF, and U&UI) exceeded the planned metric, which resulted in 8,306 actual closures with an estimated positive impact of \$94.3 million in assessments made related to these closures over plan.

In addition, the strategies (ATTI, SEP&F, and DAO) that fell below their planned closures had a higher average dollar per return than the strategies (NRP, OFF, and U&UI) that exceeded their planned closures. The net effect was an estimated reduction in the total amount assessed of \$262.5 million (\$356.8 million in lost potential assessments less \$94.3 million from assessments over planned closures).<sup>22</sup> As shown in Figure 5, because it exceeded planned closures of revenue agent individual examinations during FY 2016 in the NRP and U&UI strategies, the SB/SE Division was unable to meet planned examination closures in the ATTI, SEP&F, and DAO strategies.

**Figure 5: Revenue Agent Individual Returns FY 2016 Impact Due to Planned Closures At Strategic Levels Not Being Met**

Strategy	Net Closures Over/(Under)	Average Dollars Per Return	Dollar Impact (Potential Lost Assessments)
NRP	3,843	\$6,882	\$26,447,526
P&RP	Met	Met	Met
OFF	910	\$17,882	\$16,272,620
ATTI	(877)	\$65,051	(\$57,049,727)
HIHW	Met	Met	Met
U&UI	3,553	\$14,513	\$51,564,689
SEP&F	(420)	\$71,117	(\$29,869,140)
OPP	Met	Met	Met
DAO	(7,992)	\$33,771	(\$269,897,832)
<b>Net Totals</b>	<b>8,306 Over/ (9,289) Under</b>	<b>N/A</b>	<b>(\$262,531,864)</b>

Source: TIGTA analysis of SB/SE Division Examination September U.S. Program Monitoring Report FY 2016.

<sup>22</sup> See Appendix IV for detailed information on the methodology used to calculate this potential impact.



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

**SB/SE Division management was aware strategies were over or under the planned closures; however, no action was taken to correct the disparity**

According to the PP&A function, the information on closures at the strategy level is diagnostic in nature. The PP&A function prepares the Briefing, which provides a high-level overview allowing SB/SE Division management to monitor the performance of the SB/SE Division Examination Program. We reviewed the FY 2016 Briefings to identify if closure trends were elevated to SB/SE Division management. While the FY 2016 Briefings identified trends and these trends were elevated to SB/SE Division management by strategy, we did not identify any documented corrective actions taken to adjust the allocation to meet the plan for each strategy in those Briefings.

According to SB/SE Division management, they had identified the disparity in the number of closures between the NRP and the other strategies. They recently reached an agreement with the RAAS, which should result in a decrease in the size of the sample returns needed for NRP examinations in the future. Expending resources to examine more returns than planned, when those returns have a lower average dollar per return, is generally not an efficient use of resources. The SB/SE Division needs to use data on the strategies to monitor the SB/SE Division Examination Program in order to maintain a balance in the types of returns examined and make the best use of its resources.

The Internal Revenue Manual (IRM) places specific emphasis on prioritizing the assignment of NRP strategic priority returns. The IRM specifies that NRP cases be assigned to examiners as quickly as possible and completed promptly. While it is possible to survey NRP returns, the IRM limits surveys before or after assignment unless certain exclusion criteria are met and must be approved by both the Area NRP Coordinator and the National NRP Coordinator. The Planning and Special Programs (PSP) function in each of the seven SB/SE Division Examination Areas is responsible for identifying, selecting, and delivering strategy returns to SB/SE Division Field Examination groups in their respective area. The monthly Briefings acknowledged that the NRP strategy closures exceeded planned closures as early as October 2015, while identifying that the U&UI strategy closures exceeded planned closures as early as May 2016. Nonetheless, revenue agents in the field started an average of 1,348 NRP and 631 U&UI individual returns during the last six months of FY 2016 based on the September FY 2016 U.S. Strategy Closures and New Starts report. By the end of FY 2016, as illustrated in Figure 1, the NRP exceeded planned closures by 3,843, and the U&UI exceeded planned closures by 3,553. Revenue agents in the field continued to examine individual NRP returns because of the priority to assign and complete NRP examinations, and the limitations on surveying NRP-selected returns.

SB/SE Division management was aware of the continued impact as indicated by the Briefings, and the Briefings did not document whether any action was taken to make appropriate corrections to divert limited field resources to other strategies that were not meeting planned closures, *e.g.*, the DAO. Without any direct targets/goals at the strategic priority level, the SB/SE Division has no requirement or incentive to take action on any known strategic-level



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

issues so long as the overall closures and planned use of resources remain within its targeted range.

### **Recommendation**

**Recommendation 1:** The Commissioner, Small Business/Self-Employed Division, should monitor performance goals and take corrective actions as necessary to achieve targets specific to staffing resources, *e.g.*, revenue agents and tax compliance officers, and establish actionable performance goals with a range for the number of closures for each strategic priority to ensure that they provide balanced coverage.

**Management's Response:** The IRS disagreed with this recommendation. In its response, the IRS stated that the SB/SE Division Examination function does not control how the examination plan is worked including the delivery of returns within many strategies, for example, taxpayer behavior drives the number of refund claims received, and the IRS can control neither the number nor the type of claims filed. The IRS added that it will continue to monitor and report closures by strategy and return category in an effort to identify causes of any shortfalls or surplus closures and take action where possible and appropriate.

**Office of Audit Comment:** IRS management controls the assignment of work, and TIGTA did not identify any factors that would prevent management from adhering to its examination plan.

Because examinations among the strategies vary in complexity and productivity, the examination plan is an important internal control to ensure that the field is conducting the right mix of examinations. We recommended that the IRS establish an acceptable range for each strategy to provide for minor adjustments or deviations. Establishing actionable performance goals with a range for the number of closures for each strategic priority will provide management additional information to help achieve the IRS's overall compliance goals and coverage objectives in the future.

The IRS also disagreed with our outcome measure, stating that our methodology assumes performance will be the same for all returns within each strategy and does not reflect other SB/SE Division Examination function goals of compliance and coverage, and that uncontrollable deviations may occur. Our outcome measure is based on actual performance and is the most accurate estimate of what the productivity would have been if the examination plan was followed.



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

**Improved Processes Have Led to Reductions in Surveyed and No-Change Examination Closures by Revenue Agents of Individual Returns for Most Strategies**

We evaluated the trends in survey and no-change examinations of individual returns by SB/SE Division revenue agents for FYs 2014 through 2016 using Table 37 reports.<sup>23</sup> Surveys involve tax returns initially selected for examination, but later closed without being examined based on actions by Area PSPs, field group managers, or revenue agents. According to SB/SE Division management, surveyed returns are not considered an examination accomplishment. The SB/SE Division Examination Field Case Selection operation does not use specific thresholds to measure the rate of surveyed returns, rather it monitors for dramatic changes in survey trends. Each strategy has to be evaluated separately, and some strategies will have higher rates of surveyed returns than other strategies.

**The SB/SE Division has reduced revenue agent survey rates of individual returns for all types of surveys**

A return selected for examination can be disposed by survey, if certain criteria are met. Surveys are initiated by the group manager or examiner if there has not been any contact with the taxpayer or the taxpayer's representative. Surveys are allowed by the PSP when the return has not been assigned to an examination group and there is not enough time on the statute to conduct an examination. In the SB/SE Division, individual returns can be surveyed using several processes.

- **Surveys Before Assignment by the PSP or Group** – A return selected for examination is considered surveyed before assignment if it is disposed of without contact with taxpayers, or their representatives, and before assignment to an examiner. Surveys before assignment for revenue agent field groups are typically determined and prepared by the group manager.
- **Surveys After Assignment** – A return shall be reported as a survey after assignment if the examiner, after consideration of the return and without contact with taxpayers or their representatives, believes that an examination of the return would result in no material change in tax liability. According to SB/SE Division management, a survey after assignment is prepared by the revenue agent using Form 1900, *Income Tax Survey*, after reviewing the return for audit potential. The group manager then approves the survey.
- **Error Accounts With No Returns** – Used to remove records from the Audit Information Management System (AIMS) database that were established in error. According to SB/SE Division management, one common reason for using error accounts

---

<sup>23</sup> Table 37 reports are examination-monitoring reports based on the Audit Information Management System and provide examination accomplishments.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

with no returns is for Offshore Voluntary Disclosure Initiative examinations. The error accounts with no returns process is used to eliminate entities not responsible for the Offshore Voluntary Disclosure during the examination as the agent determines which entity is responsible for the Offshore Voluntary Disclosure.

- **Surveyed Excess Inventory** – The return has audit potential but time prohibits starting the examination. Surveying excess inventory is usually conducted by each PSP Area to remove older unassigned inventory.

The largest number of surveys are Excess Inventory. According to SB/SE Division management, Excess Inventory surveys are mainly returns surveyed by PSP Area Offices and are typically not closed as a survey based on decisions by group managers or the revenue agents. Figure 6 provides a summary of SB/SE Division revenue agent individual return closures and surveys by type for FYs 2014 through 2016. Table 37 data indicate all types of surveys have declined since FY 2014, while the percentage of total surveys to individual return closures and surveys has also declined.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 6: SB/SE Division Revenue Agent Individual Return Surveys by Type for FYs 2014 Through 2016**

Type of Survey	FY 2014	FY 2015	FY 2016
Excess Inventory	23,743	20,913	9,485
Survey Before Assignment	5,928	4,022	2,746
Survey After Assignment	6,731	4,975	3,520
Error Accounts With No Returns	6,470	2,106	1,862
Other <sup>24</sup>	991	880	622
<b>Total Individual Returns Surveyed</b>	<b>43,863</b>	<b>32,896</b>	<b>18,235</b>
Total Examined Individual Returns Closed <sup>25</sup>	126,273	114,923	102,918
Total Examined Individual Returns Closed and Surveyed	170,136	147,819	121,153
<b>Percentage of Individual Returns Surveyed</b>	<b>26%</b>	<b>22%</b>	<b>15%</b>

Source: Table 37 Data Center Warehouse reports and TIGTA analysis.

**SB/SE Division survey rates for individual returns surveyed by revenue agents have decreased for most strategies**

TIGTA’s review of IRS Table 37 data indicates a significant downward trend in the number of individual returns surveyed from 43,863 in FY 2014 to 18,235 in FY 2016. TIGTA also noted the percentage of total individual returns surveyed to the total of examined individual returns closed and individual returns surveyed has also decreased each year, and that the downward trend in the number of individual returns surveyed is not just due to lower examination activity.

Our analysis of IRS Table 37 data for surveys by strategy indicates most strategy survey levels have decreased from FYs 2014 to 2016. Figure 7 shows examined individual returns closed and individual return survey levels by strategy for FYs 2014 and 2016. Overall, surveys significantly decreased from FYs 2014 to 2016.

<sup>24</sup> Other includes Table 37 columns titled “Ref Info” and “Returns.”

<sup>25</sup> Total Examined Individual Returns Closed does not include Total Individual Return Surveys.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 7: Revenue Agent Examined Individual Returns Closed and Surveyed by Strategy for FYs 2014 and 2016<sup>26</sup>**

Strategy	FY 2014				FY 2016			
	Examined Returns Closed	Returns Surveyed	Total Closed and Surveyed	Percentage Surveyed	Examined Returns Closed	Returns Surveyed	Total Closed and Surveyed	Percentage Surveyed
NRP	10,127	157	10,284	2%	16,894	155	17,049	1%
P&RP	5,257	616	5,873	10%	2,600	178	2,778	6%
OFF	16,979	2,131	19,110	11%	17,059	1,587	18,646	9%
ATTI	3,181	645	3,826	17%	2,227	173	2,400	7%
HIHW	5,742	608	6,350	10%	4,901	347	5,248	7%
U&UI	4,709	275	4,984	6%	9,795	906	10,701	8%
SEP&F	1,986	274	2,260	12%	1,692	732	2,424	30%
OPP	15,841	2,544	18,385	14%	14,654	2,114	16,768	13%
DIF	48,878	34,009	82,887	41%	26,024	11,212	37,236	30%
All Other	13,573	2,604	16,177	16%	7,072	831	7,903	11%
<b>Totals</b>	<b>126,273</b>	<b>43,863</b>	<b>170,136</b>	<b>26%</b>	<b>102,918</b>	<b>18,235</b>	<b>121,153</b>	<b>15%</b>

Source: Table 37 Data Center Warehouse reports and TIGTA analysis.

The strategic priorities also influence the percentage of surveys. For example, the NRP has the lowest survey percentage (1 percent for FY 2016), and based on the IRM, surveys of NRP returns are limited.<sup>27</sup> Further, the results of NRP examinations assist the IRS in meeting its obligations under the Improper Payments Information Act of 2002 and subsequent legislation and guidance for annual reporting on improper payments of refundable tax credits.<sup>28</sup> For these reasons, the NRP is provided as the highest strategic priority. We previously identified that the DIF is built into the examination plan after allocations have been made to the other strategies. Therefore, the DIF is considered the lowest strategic priority. TIGTA noted the possibility for increases in high DIF-scored returns selected for examination being disposed as Excess Inventory when strategic priorities such as the NRP and the U&UI exceed the planned examinations. For example, 8,151 of the 9,485 returns disposed of as Excess Inventory in FY 2016 from Figure 6 were DIF returns. As a result, the DIF generally has the highest survey percentage (30 percent for FY 2016). TIGTA also identified upward trends for individual return surveys for the U&UI and SEP&F strategies from FYs 2014 to 2016. TIGTA noted that the U&UI and SEP&F programs were the only strategic priorities with increased surveys in

<sup>26</sup> Surveyed returns include all survey types from Figure 6.

<sup>27</sup> IRM 4.22.4.2.3(1) and 4.22.4.3(1) (June 6, 2018).

<sup>28</sup> Pub. L. No. 107-300, 116 Stat. 2350.



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

proportion to examined returns since FY 2014. We noted the percentage of total individual surveys to the total of examined individual returns closed and surveyed individual returns has decreased since FY 2014.

As part of this review, we noted the number of examined U&UI returns closed increased from FY 2014 to FY 2016 from 4,709 to 9,795, while the survey percentage increased from 6 percent to 8 percent. According to SB/SE Division management, the Information Return Document Matching program, which is a large part of the U&UI strategy, has been paused and is currently in the process of being evaluated.

SB/SE Division management stated the increase in surveys from FYs 2014 to 2016 in the SEP&F strategy is a consequence of an increase in surveys of Criminal Restitution cases. According to SB/SE Division management, there were 26 Criminal Restitution surveys in FY 2014, 353 in FY 2015, and 469 in FY 2016. In March 2014, the SB/SE Division Criminal Restitution Pilot became effective following legislation allowing the IRS to assess as tax a restitution order that arises from a criminal conviction involving a defendant's failure to pay Federal taxes.<sup>29</sup> According to SB/SE Division management, the restitution ordered by the district court is assessed first and then the SB/SE Division considers related tax returns for audit potential.<sup>30</sup> SB/SE Division management noted Criminal Restitution cases are sent to Field Examination and reviewed and then surveyed if a determination is made not to examine the return based on normal survey procedures. Prior to the law change in August 2010, the IRS had no authority to assess the court-ordered restitution.

According to discussions with the SB/SE Division, the recommended inventory for each group is two and one-half months of unstarted inventory. In the last two to three years, PSP Areas have implemented new return ordering standards and programs to monitor inventory levels closely. According to the IRS, the SB/SE Division Examination Return Selection function conducted training on September 30, 2014, on the use of Control-D.<sup>31</sup> The Examination Return Selection function also updated the DIF return ordering standards as of June 2014.

TIGTA noted that the SB/SE Division has reduced all types of surveys of individual returns from FYs 2014 through 2016. This reduction appears to be the result of improved processes aimed at streamlining the inventory ordering process and decreasing the amount of inventory being held by Field Examination groups.

---

<sup>29</sup> Firearms Excise Tax Improvement Act of 2010, Pub. L. No. 111-237, 124 Stat 2497.

<sup>30</sup> According to the IRS, it is rare that there is no restitution assessed. However, when there is a survey and no restitution has been assessed, the most common reasons for surveying are the lack of availability of records and doubt as to collectability.

<sup>31</sup> Control D monitors inventory with the goal of getting the best work possible.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

**SB/SE Division revenue agents have reduced no-change examinations of individual returns for most strategies**

“No-change” examinations are examined returns closed with no adjustments or changes in liability for tax, penalties, or refundable credits. The IRS tries to reduce noncompliance by selecting returns where examiners are likely to find areas of noncompliance and recommend changes to one or more items reported on the return.

Our review of IRS Table 37 data indicates the percentage of individual return examinations closed as a no-change has decreased over the last three years from 10 percent for FY 2014 to 8 percent for FY 2016. Overall, examinations closed as a no-change have decreased from roughly 12,050 for FY 2014 to 8,274 for FY 2016. Statistics indicate that the SB/SE Division has been successful in reducing no-change examinations by SB/SE Division revenue agents overall for FYs 2014 through 2016.

TIGTA’s review of IRS Table 37 data for no-change rates by strategic priority indicates most strategic priority no-change rates have either slightly decreased or remained steady since FY 2014. Figure 8 shows no-change percentages and returns closed by priority for individual examinations for FYs 2014 and 2016.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 8: Revenue Agent Examined Individual Returns Closed As No-Change by Strategy for FYs 2014 and 2016<sup>32</sup>**

Strategy	FY 2014			FY 2016		
	Examined Returns Closed	Examined Returns No-Changed	% No-Change	Examined Returns Closed	Examined Returns No-Changed	% No-Change
NRP	10,127	1,114	11%	16,894	2,196	13%
P&RP	5,257	263	5%	2,600	104	4%
OFF	16,979	170	1%	17,059	171	1%
ATTI	3,181	286	9%	2,227	111	5%
HIHW	5,742	632	11%	4,901	196	4%
U&UI	4,709	471	10%	9,795	1,077	11%
SEP&F	1,986	99	5%	1,692	68	4%
OPP	15,841	950	6%	14,654	733	5%
DIF	48,878	6,843	14%	26,024	3,123	12%
All Other	13,573	1,222	9%	7,072	495	7%
<b>Totals</b>	<b>126,273</b>	<b>12,050</b>	<b>10%</b>	<b>102,918</b>	<b>8,274</b>	<b>8%</b>

Source: Table 37 Data Center Warehouse reports and TIGTA analysis.

We also prepared a trend analysis of individual return no-change rates by Area and strategic priority for FYs 2014 through 2016. Through this analysis, TIGTA identified one significant upward trend for no-change examinations for the Gulf States Area under the OFF strategy. The Gulf States had no-change rates for the OFF strategy of 17 percent for FY 2014, 16 percent for FY 2015, and 32 percent for FY 2016, while the overall no-change rate for OFF examinations was 1 percent for all three years. SB/SE Division management noted a decision was made to have the Gulf States remotely work “Quiet Disclosure”<sup>33</sup> OFF cases and that the Gulf States no-change rate for these types of returns was not high in comparison to other areas.

The SB/SE Division has been successful in minimizing both surveys and no-change examinations of individual returns by revenue agents overall and across most strategic priorities and areas. According to SB/SE Division management, the decrease in surveys and no-change examinations appears to be a consequence of changes in the return ordering process and new

<sup>32</sup> The Percentage No Changed column percentages are based on the percent of examined returns resulting in no changes as reported on Table 37. The Returns No Changed column is computed by multiplying the Percentage No Changed column by Examined Returns Closed column and is not directly reported on Table 37.

<sup>33</sup> According to SB/SE Division management, a “quiet disclosure” occurs when a taxpayer files an amended return reporting income from offshore bank accounts.



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

tools and techniques emphasizing ordering returns with the best audit potential. The emphasis on reducing surveys and no-change examinations should continue through detailed review of trends by strategy and implementing changes when needed.

### **There Are Opportunities to Improve Discriminant Function Scoring of Individual Returns**

The DIF is a mathematical technique used to computer score income tax returns for examination potential for the return as a whole. Examination potential is indicated by a numeric score assigned to each return by examination class. The higher the score, the greater the examination potential within each examination class. The returns are categorized into examination class by the amount of income or assets.<sup>34</sup>

As previously stated, the PSP function in each of the seven SB/SE Division Examination function Areas is responsible for identifying, selecting, and delivering strategy returns, including DIF returns to the SB/SE Division Field Examination groups in its respective area.<sup>35</sup> Returns are identified in descending DIF score order for classification (screening).<sup>36</sup> Classification is the process of determining whether a return should be selected for audit, the initial issues to be audited, and who should conduct the audit.<sup>37</sup> Experienced examiners (tax compliance officers, tax auditors, or revenue agents) detailed to the PSP function classify the returns.<sup>38</sup> Returns classified in the Compliance Data Environment are maintained in virtual inventory until the Examination function requests additional inventory to work, at which time the returns are delivered in descending DIF score order. The Compliance Data Environment prints the Compliance Data Environment facsimile of the tax return and related information.<sup>39</sup> An employee in the PSP function assembles the documents and delivers them to the appropriate Examination function group manager.<sup>40</sup>

The IRM states that generally the higher the DIF score, the greater the probability of significant tax change. The DIF score for returns of different examination classes are not comparable.<sup>41</sup> For example, one examination class may include individual income tax returns containing a Schedule C, *Profit or Loss From Business*, or Schedule F, *Profit or Loss From Farming*, while

---

<sup>34</sup> IRM 4.1.1.1.6(11) and (16) (Oct. 25, 2017).

<sup>35</sup> IRM 4.1.5.1.1(1) (Oct. 20, 2017).

<sup>36</sup> IRM 4.1.5.3.1(4) (Oct. 20, 2017).

<sup>37</sup> IRM 4.1.5.3(2) (Oct. 20, 2017).

<sup>38</sup> National SB/SE Classification Guidelines (page 1) - Classifiers are selected because of their technical expertise, knowledge, and experience as a tax compliance officer, tax auditor, or revenue agent.

<sup>39</sup> An information return, *i.e.*, Schedule K-1, *Partner's Share of Income, Deductions, Credits, etc.*, is a mandatory tax document that individuals or entities must use to notify the IRS about certain transactions.

<sup>40</sup> In 2017, the SB/SE Division Examination Central Area conducted a pilot study in which returns were delivered directly to the group from the Compliance Data Environment electronically.

<sup>41</sup> IRM 4.19.11.2.2(2) (Oct. 11, 2017).



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

another examination class may include individual income tax returns without a Schedule C or Schedule F. It is necessary to manually screen most returns received to identify the issues in need of examination and to eliminate those returns that do not warrant examination.<sup>42</sup> Although the DIF score indicates the overall tax change potential of the return as a whole, the IRS cautions the screeners that significant items that the classifier identifies may or may not have been scored, just as the absence of any item(s) may be equally important in the scoring process.<sup>43</sup>

According to the RAAS, the DIF score predicts the likelihood of a significant tax change based on variables that include transcribed tax return lines, as well as ratios between line entries and combinations of line entries.<sup>44</sup> The DIF is not designed to identify specific issues of noncompliance. Rather, it identifies the relative potential for overall noncompliance, without explaining the reason why there is a significant tax change. Returns should be selected from the highest score indicating the likelihood that an audit will result in an additional tax assessment. The DIF score does not take into account the administrative cost that would be associated with the examination.

Beginning in FY 2016, the SB/SE Division's examination plan has two major components: number of return closures and number of return starts.<sup>45</sup> In implementing the plan, the SB/SE Division strategy is to have coverage across return categories that fall within the purview of the SB/SE Division. Prior to November 2014, the SB/SE Division established a national minimum DIF cutoff score, and each Examination function Area Office calculated an area DIF score for each examination class. Aside from limited exceptions, only returns with a DIF score exceeding the national minimum DIF cutoff score were screened for examination.<sup>46</sup> Starting in November 2014, the SB/SE Division established DIF cutoff scores by SB/SE Division Examination function Area Office and examination class with no national minimum cutoff. According to SB/SE Division management, the DIF strategy approach is based on a top down approach in which the highest available DIF-scored returns are selected first for audit.

### ***For the majority of classes of revenue agent examinations, individual returns with the highest DIF scores do not result in a higher net tax assessment***

We analyzed individual returns filed with the IRS during Calendar Year 2015 and examined by revenue agents under the DIF strategy and closed during October 1, 2016, through June 30, 2017. Because they were received after November 2014, these primary returns were some of the first returns processed under the new practice in which the SB/SE Division Examination function Headquarters calculated DIF cutoff scores specific to each Examination

---

<sup>42</sup> IRM 4.19.11.2.2(2) (Oct. 11, 2017).

<sup>43</sup> IRM 4.19.11.2.4.1(2) (June 22, 2016).

<sup>44</sup> It was noted that nontranscribed lines are reviewed as well, and if one is seen as a significant indicator, the group tries to find a transcribed line that will proxy for the nontranscribed line.

<sup>45</sup> IRM 4.1.1.2 (Oct. 25, 2017).

<sup>46</sup> IRM 4.1.2.7.7(1) (Oct. 19, 2017).



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

function Area by examination class and provided them to each SB/SE Division Field Examination Area.<sup>47</sup> We evaluated the IRS's examination results for these returns by examination class and by DIF score. Within each examination class, using the assigned DIF scores, we grouped the returns into ranges and computed the average net tax change per return for each range. Depending on the actual DIF scores, the number of returns varied between ranges within an examination class. The DIF scores represented by each range are not comparable between examination classes. When computing the tax change, we included both tax assessments and abatements, hereafter referred to as net tax assessment. Range Number (No.) 1 contains returns with the lowest DIF scores, and Range No. 4 contains returns with the highest DIF scores. For each examination class, we identified the range that resulted in the highest average net tax assessment. Based on DIF-selected returns filed during Calendar Year 2015 and examined by a revenue agent during October 1, 2016, through June 30, 2017, we identified seven examination classes in which examined returns in a lower DIF score range resulted in the highest average net tax assessments among examined returns.<sup>48</sup> Figure 9 shows that, more often than not, examined returns with the highest DIF scores did not result in the highest average net tax assessment than examined returns with lower DIF scores. While this result may not be indicative of the general population of returns with the highest DIF scores, it indicates that there may be opportunities to refine and improve DIF score models.

---

<sup>47</sup> The primary return is generally the first return that an examination started with.

<sup>48</sup> The seven examination classes are identified when the highest DIF-scored returns (Range No. 4) did not result in a higher average net tax assessment as shown in Figure 9.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 9: Average Revenue Agent Examination Results Per Primary Return**

Examination Class	Number of Primary Returns	Average Net Tax Assessment Per Primary Return				
		DIF Score Range No. 1	DIF Score Range No. 2	DIF Score Range No. 3	DIF Score Range No. 4	Did the Highest DIF-Scored Returns (Range No. 4) Result in a Higher Average Net Tax Assessment?
A	36	\$1,055	<b>\$14,063</b>	\$657	\$3,722	No
B	462	**1**	\$6,825	\$19,937	<b>\$39,943</b>	Yes
C	64	**1**	\$3,163	\$6,422	\$7,621	No
D	309	\$657	\$5,168	<b>\$11,344</b>	\$10,510	No
E	291	\$7,808	\$7,200	\$8,307	<b>\$11,216</b>	Yes
F	147	\$9,080	\$9,362	<b>\$9,508</b>	\$8,986	No
G	437	**1**	<b>\$115,267</b>	\$16,997	\$16,309	No
H	25	\$541	\$1,323	<b>\$8,730</b>	(\$3,037)	No
I	341	**1**	\$7,141	\$4,615	<b>\$18,691</b>	Yes
J	814	\$19,108	\$13,269	\$8,228	<b>\$26,676</b>	Yes
K	504	\$13,978	<b>\$24,383</b>	\$4,401	\$18,400	No

Source: TIGTA analysis of closed AIMS data for FY 2017 cumulative through the third quarter.  
 \*\*\*\*\*/\*\*\*\*\*

We also computed the productivity on a “per examination hour basis” with similar results. For the majority of examination classes, the examined returns with the highest DIF scores did not result in better productivity.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 10: Audit Results Per Examination Hour for Primary Returns Examined by Revenue Agents**

Examination Class	Number of Primary Returns	Average Net Tax Assessment Per Hour				Did the Highest DIF-Scored Returns (Range No. 4) Result in a Higher Average Net Tax Assessment?
		DIF Score Range No. 1	DIF Score Range No. 2	DIF Score Range No. 3	DIF Score Range No. 4	
A	36	\$30.43	<b>\$430.51</b>	\$24.17	\$91.92	No
B	462	**1**	\$156.39	\$401.11	<b>\$655.39</b>	Yes
C	64	**1**	\$197.69	\$221.15	\$192.85	No
D	309	\$20.97	\$126.73	<b>\$278.17</b>	\$228.22	No
E	291	\$185.45	\$165.30	\$158.07	<b>\$207.94</b>	Yes
F	147	\$195.87	\$201.81	\$200.12	<b>\$208.98</b>	Yes
G	437	**1**	<b>\$3,002.61</b>	\$333.66	\$319.93	No
H	25	\$9.00	\$23.27	<b>\$168.69</b>	(\$46.68)	No
I	341	**1**	\$191.37	\$106.07	<b>\$550.89</b>	Yes
J	814	\$386.27	\$230.72	\$146.58	<b>\$725.88</b>	Yes
K	504	\$440.25	<b>\$462.29</b>	\$82.36	\$435.72	No

Source: TIGTA analysis of closed AIMS data for FY 2017 cumulative through the third quarter.

\*\*\*\*\*1\*\*\*\*\*

**Examination results by case versus return approach**

As a part of the IRS’s Future State Initiative, the SB/SE Division is refining the inventory for examinations to produce the maximum compliance impact. The Resource and Workload Mix Team developed a workload delivery model that projects output (closure volume and mix as well as productivity) based upon the mix of work started. One of the key elements of the model is the association of related returns and how they are viewed in terms of determining productivity. An examiner is normally assigned a “primary” return and, in most instances, expands from this primary return to other noncompliant returns that are “related” to the primary taxpayer. The Team discovered that a more accurate view of productivity results when considering the entire case, which includes the primary and related pick-up returns, *e.g.*, prior and subsequent returns.

In consideration of the SB/SE Division’s “case” approach to evaluate productivity, for the primary revenue agent examined individual returns that we analyzed in Figures 9 and 10, we identified the filed prior and/or subsequent year returns that were picked up by the examiner.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

We then analyzed the examination results from a “case” perspective in terms of net tax change per case and per examination hour. In evaluating the tax change amounts, we included both additional tax assessments and abatements. Figures 11 and 12 show that the highest DIF scores among examined returns do not necessarily correspond to the highest average net tax assessment.

For each examination class, we identified the DIF score range that resulted in the highest average net tax assessment. Figure 11 shows that we identified six examination classes in which a relatively lower DIF score range among the examined returns resulted in a higher average net tax assessment per case.<sup>49</sup>

**Figure 11: Average Revenue Agent Examination Results Per Case**

Examination Class	Number of Cases	Average Net Tax Assessment Per Case				Did the Highest DIF-Scored Returns (Range No. 4) Result in a Higher Average Net Tax Assessment?
		DIF Score Range No. 1	DIF Score Range No. 2	DIF Score Range No. 3	DIF Score Range No. 4	
A	36	\$1,139	<b>\$21,566</b>	\$657	\$6,132	No
B	462	**1**	\$10,587	\$31,882*	<b>\$46,064*</b>	Yes
C	64	**1**	\$5,781	\$13,451	\$11,889	No
D	309	\$657	\$8,147	\$16,958	<b>\$17,027</b>	Yes
E	291	\$7,808	\$12,578	\$11,563	<b>\$17,691</b>	Yes
F	147	\$13,429	\$12,557	\$13,650	<b>\$18,329</b>	Yes
G	437	**1**	<b>\$115,267</b>	\$25,476*	\$25,634*	No
H	25	\$541	\$1,756	<b>\$16,095</b>	(\$3,037)	No
I	341	**1**	\$10,841	\$9,315	<b>\$120,586</b>	Yes
J	814	<b>\$28,085*</b>	\$16,569	\$8,663	\$26,562*	No
K	504	\$13,978	<b>\$32,983*</b>	\$13,185*	\$18,845	No

Source: TIGTA analysis of closed AIMS data for FY 2017 cumulative through the third quarter.

\*\*\*\*\*]

\* - There are five taxpayers that had two primary returns each that were all DIF selected and fell in different Activity Codes and/or DIF score ranges. These five taxpayers also had related return pick-ups. The examination results data associated with the primary and pick-up returns are captured in these numbers.

<sup>49</sup> The six examination classes are identified when the highest DIF-scored returns (Range No. 4) did not result in a higher average net tax assessment as shown in Figure 11.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

Furthermore, as shown in Figure 12, we identified seven examination classes in which a relatively lower DIF score range among the examined returns resulted in a higher average net tax assessment per examination hour from casework.<sup>50</sup>

**Figure 12: Audit Results Per Examination Hour for Cases Examined by Revenue Agents**

Examination Class	Number of Cases	Average Net Tax Assessment Per Hour				Did the Highest DIF-Scored Returns (Range No. 4) Result in a Higher Average Net Tax Assessment?
		DIF Score Range No. 1	DIF Score Range No. 2	DIF Score Range No. 3	DIF Score Range No. 4	
A	36	\$21.76	<b>\$513.46</b>	\$21.77	\$117.12	No
B	462	**1**	\$199.34	\$511.78*	<b>\$640.60*</b>	Yes
C	64	**1**	\$260.41	\$315.15	\$200.46	No
D	309	\$18.60	\$154.87	<b>\$301.24</b>	\$248.59	No
E	291	\$185.45	\$213.09	\$159.74	<b>\$245.06</b>	Yes
F	147	\$236.21	\$227.73	\$243.39	<b>\$351.91</b>	Yes
G	437	**1**	<b>\$3,002.61</b>	\$394.29*	\$389.61*	No
H	25	\$9.00	\$26.37	<b>\$228.30</b>	(\$46.68)	No
I	341	**1**	\$252.89	\$199.18	<b>\$2,908.19</b>	Yes
J	814	<b>\$486.51*</b>	\$250.97	\$139.54	\$444.55*	No
K	504	\$440.25	<b>\$570.60*</b>	\$224.74*	\$426.48	No

Source: TIGTA analysis of closed AIMS data for FY 2017 cumulative through the third quarter.

\*\*\*\*\*]\*\*\*\*\*

\* - There are five taxpayers that had two primary returns each that were all DIF selected and fell in different Activity Codes and/or DIF score ranges. These five taxpayers also had related return pick-ups. The examination results data associated with the primary and pick-up returns are captured in these numbers.

<sup>50</sup> The seven examination classes are identified when the highest DIF-scored returns (Range No. 4) did not result in a higher average net tax assessment as shown in Figure 12.



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

These DIF formulas were developed based on available NRP data.<sup>51</sup> According to the RAAS, NRP audits are a comprehensive examination of sample returns. The DIF has provided the most closures of any single priority for several years.<sup>52</sup>

According to the RAAS, the DIF models are developed based on NRP data only. The RAAS evaluates overall DIF-selected examination results for each examination class. That result is one factor that determines which examination class's DIF model will be updated in the next DIF development cycle. Based on TIGTA's data analyses previously discussed, it may be beneficial to refine the DIF scoring model. In fact, a study prepared in August 2013 by an external contractor recommended that the IRS add enforcement history into its selection models to improve examination performance.<sup>53</sup> In an age of globalization and technological advancement, it is imperative for the IRS to capitalize on all pertinent data sources to select the most egregious returns. This would be consistent with the IRS Examination function's goal to promote the highest degree of taxpayers' voluntary compliance while making the most efficient use of finite examination staffing and other resources. Otherwise, the DIF score may not reflect the likelihood of a higher net tax assessment as shown by the results of our analysis.

When we discussed our results with RAAS and SB/SE Division management, the IRS was concerned that: 1) the time period covered by TIGTA's analysis was too narrow, 2) dividing the DIF scores into four ranges was incorrect, and 3) our results excluded examinations performed by tax compliance officers. To address these concerns, we expanded our analysis covering individual returns filed with the IRS during Calendar Year 2015. We changed the methodology so that ranges contained about the same number of returns and increased the number of DIF score ranges from four, as shown in Figure 9, to 20 as shown in Appendix V, Figure 1, and included field examinations by revenue agents and tax compliance officers through December 2017. As shown in the expanded analysis in Appendix V, Figure 1, we similarly found that, more often than not, examined individual returns by revenue agents and tax compliance officers with the highest DIF scores do not result in a higher net tax assessment for most examination classes. We identified nine examination classes in which examined returns that fell into a relatively lower DIF score range resulted in the highest average net tax assessment.<sup>54</sup>

RAAS management also stated outliers potentially skew results and therefore, examination assessments identified as outliers should be adjusted for analytical purposes. To reduce the

---

<sup>51</sup> IRM 4.1.2.7(2) (Oct. 19, 2017).

<sup>52</sup> The U.S. Program Monitoring Reports for FYs 2014 through 2016 show that revenue agents in the SB/SE Division closed more than 111,300 DIF strategy individual income tax returns. The U.S. Program Monitoring Reports also show that revenue agents only closed 51,431 in the next closest strategy (OFF) during the same period.

<sup>53</sup> Compliance Data Warehouse Body of Knowledge Template: Enforcement Revenue Information System, Version 2.0 (prepared for IRS Research, Analysis, and Statistics), August 7, 2013.

<sup>54</sup> The nine examination classes are identified when the highest DIF-scored returns (Range No. 20) did not result in a higher net tax assessment as shown in Appendix V, Figure 1.



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

impact of outliers, we performed additional analysis using the RAAS's censoring methodology.<sup>55</sup> Our analysis found that by adjusting the dollar amount of outlier returns, the average net tax assessment increased for some examination classes/ranges, decreased for others, and a few remained the same. Regardless, as shown in the expanded analysis in Appendix V, Figure 2, we similarly found that, more often than not, the highest DIF-scored individual returns examined by revenue agents and tax compliance officers do not result in a higher net tax assessment for most examination classes. We identified seven examination classes in which examined returns that fell into a relatively lower DIF score range resulted in the highest average net tax assessment.<sup>56</sup>

RAAS management also asserted that tax abatements be modified for analytical purpose by including the returns but adjusting the dollar amount to zero. The IRS believes an examination that resulted in an abatement should not penalize a selection method any more than an examination that resulted in no change in tax.<sup>57</sup> According to RAAS management, "the purpose of DIF as a workload selection method is to identify, for a given coverage level, a group of returns most likely to have significantly underreported tax." TIGTA believes that all examination results should be considered to evaluate a workload selection method, including actual tax abatements. If the tax abatements are modified by changing the dollar amounts from negative amounts to zero, the productivity is skewed to the positive and would not reflect the true compliance effect. The IRS also suggested that TIGTA should have used the approach of comparing DIF-selected examinations to returns with NRP data. However, TIGTA's analysis is focused on individual returns that were actually examined by revenue agents as selected within the DIF strategy, not those that were reviewed as part of a separate NRP strategy.

According to RAAS management, surveys are an inherent complicating factor influencing TIGTA's analysis. Some high DIF-scored returns may not have been selected during classification or may have been surveyed at the examination group level. Because surveyed returns are not examined, it is impossible to assess whether the survey process is effective at removing returns without good audit potential from the examination stream. Also, the RAAS was unaware of any analysis that has addressed this topic. Additionally, the SB/SE Division stated it does not track classification and surveyed returns by DIF score. Consequently, the IRS has no data on a survey factor that it cites as "complicates" the evaluation of returns examined under the DIF strategy.

---

<sup>55</sup> Censoring methodology in this context means that the examination results for the top 2 percent of recommended net tax assessments is set to equal to the 98<sup>th</sup> percentile, and the bottom 2 percent is set to equal to the 2<sup>nd</sup> percentile. The replacement low and high dollar amounts are computed separately for each examination class. After censoring, it is possible that a replacement value be a negative (abatement).

<sup>56</sup> The seven examination classes are identified when the highest DIF-scored returns (Range No. 20) did not result in a higher net tax assessment as shown in Appendix V, Figure 2.

<sup>57</sup> The abatement exclusion assertion is independent and separate from the censoring issue. The RAAS believes it is possible that these examinations resulted in a smaller abatement than had the taxpayer submitted an amended return refund claim absent the examination.



## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

Notwithstanding the IRS's various contentions, the assigned DIF score largely dictates what happens to the tax return—it influences selection, and priority for screening and examination. A DIF score that does not fully take into consideration compliance risks may exclude an underreported tax return from being examined and/or result in examining a generally compliant return. As TIGTA demonstrated based on analyses of returns received during Calendar Year 2015 and examined between October 1, 2016, and June 30, 2017, the examined returns with the highest DIF scores do not necessarily result in the highest net examination adjustment. While TIGTA agrees with the RAAS that no method can perfectly predict underreported tax, our results indicate that there are opportunities to refine the DIF scoring, especially for certain examination classes.

In discussing the audit results with the IRS, RAAS management expressed a concern that TIGTA only looked at returns in the top 2 to 3 percent of DIF-scored returns to draw conclusions about the overall effectiveness of the DIF. However, when selecting returns as part of its DIF examination strategy, the IRS predominantly selects returns in the top 2 to 3 percent of DIF-scored returns. The RAAS also asserted that TIGTA did not compare the DIF-selected returns to a representative sample of returns not selected by the DIF. TIGTA used this approach because the IRS considers DIF-selected returns to have high potential for a significant tax change.<sup>58</sup> As previously stated, the IRS's DIF examination strategy consists of returns that were actually selected for examination and does not include all returns assigned a DIF score. TIGTA's conclusions are based on examined returns representing the majority of DIF strategy selected returns.

While TIGTA considered the concerns expressed by the IRS, these concerns do not explain why, more often than not, the highest DIF scores for individual returns examined by revenue agents and tax compliance offers did not result in higher net tax assessments than lower DIF scores for most examination classes.

---

<sup>58</sup> Our audit objective and scope focused on individual returns selected for examination. As such, we did not validate the RAAS's statement that examined returns come from the top 2 to 3 percent of DIF scores for filed returns in each examination class, which includes non-examined returns outside the scope of this review.



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

### **Recommendation**

**Recommendation 2:** The Chief Research and Analytics Officer should evaluate the examination results to determine whether refinements could improve the correlation of DIF scores and net tax assessments for individual returns examined by revenue agents and tax compliance officers.

**Management's Response:** The IRS disagreed with this recommendation. The IRS stated that it regularly updates the DIF scoring models based on NRP results, striving to improve upon the already effective models. The IRS further stated that in its routine reviews of DIF scoring models and independent of this audit, 10 of the 11 examination classes that were reviewed in the report have already been updated and implemented since the processing year that was evaluated. Therefore, the IRS's position is that the report's specific results do not provide any actionable information relevant to the current DIF models. Although the IRS remains committed to continuing its efforts to evaluate DIF performance and improve the performance of the DIF scoring models, it does not agree that the analysis in this report supports any additional corrective action.

**Office of Audit Comment:** In this report, TIGTA demonstrated that, more often than not, higher DIF scores do not yield higher adjustments by revenue agents on individual tax returns. The IRS states that NRP data validates the current DIF models; however, the IRS derives the DIF models and adjusts the DIF models from NRP data. The IRS uses the NRP as the primary mechanism for developing and refining the DIF, and evaluating examination results outside of the NRP would help further refine the DIF to improve the correlation between the DIF score and examination results.



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

## **Appendix I**

### *Detailed Objective, Scope, and Methodology*

The objective of this review was to perform trend analyses of the SB/SE Division's strategic priority selection methods used to identify individual income tax returns for examination by revenue agents<sup>1</sup> and assess the effectiveness of the strategic priorities with an emphasis on the DIF selection method. To accomplish our objective, we:

- I. Determined whether the SB/SE Division established performance goals to assess the effectiveness of each strategic priority with an emphasis on evaluating the DIF selection priority method used to identify individual income tax returns for examination.
  - A. Obtained the U.S. Program Monitoring reports and the U.S. Monthly Performance reports for FYs 2014 through 2016.
  - B. Reviewed the U.S. Program Monitoring reports and the U.S. Monthly Performance reports for FYs 2014 through 2016 to determine whether the SB/SE Division established goals to assess the effectiveness of each strategic priority.
  - C. Obtained and analyzed the Briefings for evidence that trend information was elevated to management in the SB/SE Division Examination function.
  - D. Interviewed SB/SE Division management to discuss the definition and goal of each strategic priority and to determine whether the SB/SE Division established performance goals to assess the effectiveness of each strategic priority.
  - E. Obtained and reviewed Strategic Priority Coding Lists for FYs 2014 through 2016 to determine which strategies the SB/SE Division used between FYs 2014 and 2016.
- II. Evaluated the effectiveness of the strategic priorities used to identify individual income tax returns for examination by revenue agents.
  - A. Prepared a trend analysis, using data from the U.S. Program Monitoring reports for FYs 2014 through 2016, comparing selected metrics for each strategic priority.
- III. Determined how the SB/SE Division monitors and controls surveys and examinations closed as no changes through the examination case selection process.
  - A. Prepared a trend analysis using Table 37 reports for FYs 2014 through 2016 to analyze the survey rates by type and priority across the seven examination areas. We discussed our analysis with the IRS.

---

<sup>1</sup> See Appendix VI for a glossary of terms.



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

- B. Prepared a trend analysis using Table 37 reports for FYs 2014 through 2016 to analyze the no-change rate by priority across the seven examination areas. We discussed our analysis with the IRS.
  - C. Validated the completeness and accuracy of the Table 37 reports for both surveys and no-change data. We obtained AIMS data for non-examined and examined closures using the TIGTA Data Center Warehouse for FY 2014, FY 2015, and FY 2016. Prior to use, we verified the completeness of non-examined closures and examined closures against the Table 37 reports. Additionally, we verified the accuracy of the records against the IRS's Integrated Data Retrieval System for FYs 2014 through 2016.
- IV. Determined if higher DIF scores resulted in more significant net tax change than lower DIF scores.
- A. Obtained from RAAS management information regarding the DIF model and what the DIF score represents and predicts.
  - B. Obtained and reviewed from SB/SE Division management documentation regarding its DIF strategy, examination coverage, and DIF cutoff scores.
  - C. Researched the IRM and other pertinent IRS documentation associated with the DIF strategy and scoring.
  - D. Analyzed AIMS data on primary individual income tax return examinations closed by the SB/SE Division during the period October 2016 through June 2017. We further limited the returns to those that met the following criteria, which resulted in 3,430 primary returns.
    - 1. The IRS received the return during Calendar Year 2015.
    - 2. The return was selected as part of the DIF strategy.
    - 3. The examination was conducted by revenue agents located in Area Offices.
    - 4. The return was categorized in an examination class with a DIF cutoff score.
  - E. For each of the 3,430 primary returns identified in Step IV.D., analyzed the net tax change and the examination hours incurred. Thereafter, we summarized the results by examination class by DIF score. To determine if the higher DIF scores resulted in higher net tax change within each examination class, we separated the examination results by DIF score into four ranges. We computed the average net tax change per return and per examination hour.
  - F. For the 3,430 primary returns identified in Step IV.D., identified the associated filed prior and/or subsequent year returns that were picked up by the examiner as part of the case review. We analyzed the net tax change and total examination hours incurred for each case. Thereafter, we summarized the case results by examination



---

## *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

- class by DIF score. To determine if the higher DIF scores resulted in higher net tax change within each examination class, we separated the examination results by DIF score into four ranges for comparison. We computed the average net tax change per return and per examination hour.
- G. Reviewed the RAAS-provided examination data on individual returns filed with the IRS during Calendar Year 2015 and examined under the DIF strategy for Field Examination through December 2017. We separately analyzed the actual and RAAS-censored net tax change. To determine if the higher DIF scores resulted in a higher tax change within each examination class, we segregated the returns into 20 ranges based on return volume and DIF score. Each range contained about the same number of returns. We computed the average actual and censored net tax assessment for each range for comparison. Range No. 1 contained returns with the lowest DIF scores, and Range No. 20 contained returns with the highest DIF scores.
- H. Validated the source relied on in IRS examination results data analyses. We obtained data from the IRS's closed AIMS data for FY 2017 cumulative through the third quarter that were available on TIGTA's Data Center Warehouse. Prior to use, we verified the completeness of record count against the Table 37 report. Additionally, we verified the accuracy of 10 taxpayers' records against the IRS's Integrated Data Retrieval System and the Correspondence Examination Automation Support System.

### **Internal controls methodology**

Internal controls relate to management's plans, methods, and procedures used to meet their mission, goals, and objectives. Internal controls include the processes and procedures for planning, organizing, directing, and controlling program operations. They include the systems for measuring, reporting, and monitoring program performance. We determined that the following internal controls were relevant to our audit objective: the IRS's policies, procedures, and practices related to the selection of individual income tax returns for examination by revenue agents. We evaluated these controls by contacting management, reviewing IRM guidance provided to managers and employees, reviewing reports, and analyzing closed examination data.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

---

**Appendix II**

*Major Contributors to This Report*

Matthew A. Weir, Assistant Inspector General for Audit (Compliance and Enforcement Operations)  
Christina M. Dreyer, Director  
Javier Fernandez, Audit Manager  
David Hartman, Lead Auditor  
Lee Hoyt, Senior Auditor  
Julia Tai, Senior Auditor



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

**Appendix III**

*Report Distribution List*

Commissioner, Small Business/Self-Employed Division  
Chief Research and Analytics Officer  
Deputy Commissioner for Operations Support  
Deputy Commissioner for Services and Enforcement  
Deputy Commissioner, Small Business/Self-Employed Division  
Director, Examination, Small Business/Self-Employed Division  
Director, Knowledge Development and Application  
Director, Office of Audit Coordination



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

## Appendix IV

### Outcome Measure

This appendix presents detailed information on the measurable impact that our recommended corrective actions will have on tax administration. This benefit will be incorporated into our Semiannual Report to Congress.

#### **Type and Value of Outcome Measure:**

- Increased Revenue – Potential; 17,595 taxpayer accounts, netting \$262.5 million for FY 2016 (see page 7).

#### **Methodology Used to Measure the Reported Benefit:**

In our review of the SB/SE Division Examination September 2016 U.S. Program Monitoring Report, we found that three strategies (NRP, OFF, and U&UI) exceeded planned closures and four strategies (ATTI, SEP&F, and DAO combined) did not meet planned closures while the Revenue Agent-Total Closures met its target (99 percent) for FY 2016.

The SB/SE Division has established planned performance targets/goals for combined revenue agent<sup>1</sup> and tax compliance officer closures; however, it does not have any performance goals that it must meet at the strategic priority level. According to SB/SE Division management, they monitor and measure program performance for diagnostic purposes only and do not require a response; *i.e.*, there is no required action for strategic priorities that are performing over or under planned metrics. While there are no performance goals at the strategic level, the IRM places specific emphasis on the assignment of NRP strategic priority returns and limits surveys before or after assignment. The GAO's *Standards for Internal Control in the Federal Government* provides under the Monitoring internal control that management should establish and operate monitoring activities to assess the quality of performance and promptly take corrective actions in order to achieve objectives.<sup>2</sup>

TIGTA calculated the outcome measure over the population of closed individual return examinations by revenue agents for FY 2016. The calculations are based on metrics included in the September 2016 U.S. Program Monitoring Report, as listed for revenue agent closures of individual return examinations.<sup>3</sup> Figure 1 details how TIGTA arrived at the overall outcome measure of \$262.5 million.

---

<sup>1</sup> See Appendix VI for a glossary of terms.

<sup>2</sup> GAO, GAO-14-704G, *Standards for Internal Control in the Federal Government* (Sept. 2014).

<sup>3</sup> Fiscal year cumulative data in the U.S. Program Monitoring Reports are as of the month end and may differ from final AIMS data.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 1: Outcome Measure Calculation of Revenue Agent Individual Returns  
FY 2016 Impact Due to Planned Closures At Strategic Levels Not Being Met**

	A	B	C	D	E	F	G
Strategy	Actual Closures	Planned Closures	% of Plan (A/B)	Plan Met C < 95% = Under C 95% to 105% = Met C > 105% = Over	Net Returns Over or (Under) (A-B)	Dollars Per Return	Dollar Impact (Potential Lost Assessments) (E x F)
NRP	16,904	13,061	129%	Over	3,843	\$6,882	\$26,447,526
P&RP	2,601	2,609	100%	Met	Met	Met	Met
OFF	17,059	16,149	106%	Over	910	\$17,882	\$16,272,620
ATTI	2,227	3,104	72%	Under	(877)	\$65,051	(\$57,049,727)
HIHW	4,908	4,855	101%	Met	Met	Met	Met
U&UI	9,802	6,249	157%	Over	3,553	\$14,513	\$51,564,689
SEP&F	1,682	2,102	80%	Under	(420)	\$71,117	(\$29,869,140)
OPP	14,651	14,693	100%	Met	Met	Met	Met
DAO	32,509	40,501	80%	Under	(7,992)	\$33,771	(\$269,897,832)
<b>Net/Totals</b>	<b>102,343</b>	<b>103,323</b>	<b>99%</b>	<b>Met Overall Plan for Strategic Priorities</b>	<b>17,595 Impacted Closures<sup>4</sup></b>	<b>N/A</b>	<b>(\$262,531,864) Net Potential Lost Assessments</b>

Source: TIGTA analysis of SB/SE Division Examination September U.S. Program Monitoring Report FY 2016.

<sup>4</sup> The 17,595 closures include both the 8,306 closures for the strategies that were over accomplished and the 9,289 closures for the strategies that were under accomplished.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Appendix V**

*Results of Individual Primary Return Examinations by Revenue Agents and Tax Compliance Officers*

**Figure 1: Average Field Examination Results Per Primary Return**

Examination Class	Examination Class/Average Net Tax Assessment Per Primary Return										
	A	B	C	D	E	F	G	H	I	J	K
Number of Primary Returns	1,343	4,231	6,226	9,771	7,431	3,862	2,571	115	1,418	2,852	1,148
DIF Score Range No. 1	<b>\$3,610</b>	\$17,656	\$3,421	\$3,921	\$5,396	\$9,191	\$21,846	\$10,662	\$7,546	\$12,201	\$18,131
DIF Score Range No. 2	\$2,514	\$16,526	\$3,397	\$4,211	\$5,619	\$10,196	\$16,138	\$838	\$2,938	\$14,611	\$9,850
DIF Score Range No. 3	\$3,018	\$19,895	\$3,844	\$5,776	\$5,125	\$9,726	\$9,659	\$402	\$5,474	\$13,933	\$14,839
DIF Score Range No. 4	\$2,607	\$18,849	\$4,174	\$5,457	\$5,424	\$10,287	\$18,455	\$7,393	\$9,411	\$10,340	\$6,795
DIF Score Range No. 5	\$3,238	\$13,876	\$4,320	\$6,038	\$5,485	\$9,976	\$21,706	\$1,004	\$5,037	\$16,162	\$18,519
DIF Score Range No. 6	\$2,461	\$18,425	\$4,446	\$5,891	\$6,146	\$9,308	\$18,118	\$3,637	\$5,318	\$14,014	\$55,531
DIF Score Range No. 7	\$3,125	\$20,631	\$5,568	\$5,892	\$5,409	\$9,877	\$17,522	\$68	\$8,389	\$23,253	\$20,668
DIF Score Range No. 8	\$2,118	\$20,827	\$5,623	\$6,063	\$6,714	\$10,096	\$13,425	<b>\$19,384</b>	\$5,977	\$20,406	\$10,908
DIF Score Range No. 9	\$2,654	\$21,479	\$5,116	\$6,878	\$6,404	\$8,982	\$21,971	\$436	\$3,288	\$15,938	\$9,709
DIF Score Range No. 10	\$2,480	\$19,110	\$5,610	\$6,601	\$5,731	\$10,047	\$19,738	\$538	\$10,545	<b>\$26,697</b>	\$6,849
DIF Score Range No. 11	\$2,501	\$20,348	\$6,129	\$7,199	\$5,218	\$12,069	\$21,780	\$2,272	\$8,943	\$21,481	\$57,625
DIF Score Range No. 12	\$2,864	\$19,828	\$5,799	\$7,196	\$6,456	\$10,989	\$26,002	\$135	\$8,184	\$20,716	\$30,316
DIF Score Range No. 13	\$2,291	\$23,196	\$6,699	\$7,636	\$6,434	\$11,065	\$20,844	\$4,043	\$2,526	\$15,283	\$25,764
DIF Score Range No. 14	\$2,605	\$16,751	\$6,739	\$7,258	\$6,932	\$10,253	\$22,230	\$761	\$8,273	\$18,070	\$22,845
DIF Score Range No. 15	\$2,936	\$20,982	<b>\$7,426</b>	\$8,615	\$7,468	\$11,245	<b>\$26,011</b>	\$710	\$5,218	\$22,517	<b>(\$34,452)</b>
DIF Score Range No. 16	\$2,778	\$12,910	\$6,766	\$9,068	\$7,260	\$12,203	\$19,060	\$5,291	<b>\$11,677</b>	\$18,289	<b>(\$6,377)</b>
DIF Score Range No. 17	\$3,014	\$29,736	\$7,265	\$9,418	\$7,307	\$12,082	\$18,658	\$425	\$5,932	\$14,893	\$16,249
DIF Score Range No. 18	\$2,131	<b>\$36,213</b>	\$7,203	\$10,542	\$7,735	\$11,843	\$22,546	<b>(\$2,363)</b>	\$8,106	\$22,336	\$10,901
DIF Score Range No. 19	\$2,964	\$24,643	\$6,291	\$10,018	\$8,746	<b>\$12,303</b>	\$23,177	\$12,298	\$7,524	\$11,516	<b>\$62,290</b>
DIF Score Range No. 20	\$2,933	\$25,671	\$7,200	<b>\$11,151</b>	<b>\$9,377</b>	\$10,771	\$19,684	\$2,307	\$8,681	\$21,641	\$39,255
Did DIF Score Range No. 20 Result in Higher Average Net Tax Assessment?	No	No	No	Yes	Yes	No	No	No	No	No	No

Source: TIGTA analysis of RAAS-provided Calendar Year 2015 DIF-selected primary return examinations closed through December 2017.

For each examination class, we segregated the data into 20 ranges based on return volume and DIF score. Each range contained about the same number of returns. We computed the average



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

net tax assessment for each range for comparison. Range No. 1 contains returns with the lowest DIF scores among examined returns, and Range No. 20 contains returns with the highest DIF scores. As Figure 1 shows, even with inclusion of tax compliance officers and a wider time frame, more often than not, the examined returns with the highest DIF scores still did not result in a higher average net tax assessment than examined returns with lower DIF scores.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Figure 2: Average Censored Field Examination Results Per Primary Return<sup>1</sup>**

Examination Class	Examination Class/Average Censored Net Tax Assessment Per Primary Return										
	A	B	C	D	E	F	G	H	I	J	K
<b>No. of Primary Returns</b>	1,343	4,231	6,226	9,771	7,431	3,862	2,571	115	1,418	2,852	1,148
<b>DIF Score Range No. 1 Difference<sup>2</sup></b>	\$2,641 (\$969)	\$17,171 (\$485)	\$3,425 \$4	\$3,950 \$29	\$4,998 (\$398)	\$8,614 (\$577)	\$17,602 (\$4,244)	<b>\$10,662</b> \$0	\$7,945 \$399	\$10,629 (\$1,572)	\$20,456 \$2,325
<b>DIF Score Range No. 2 Difference</b>	\$2,390 (\$124)	\$16,360 (\$166)	\$3,349 (\$48)	\$4,167 (\$44)	\$5,286 (\$333)	\$9,646 (\$550)	\$15,837 (\$301)	\$838 \$0	\$2,938 \$0	\$13,702 (\$909)	\$9,850 \$0
<b>DIF Score Range No. 3 Difference</b>	\$2,797 (\$221)	\$17,093 (\$2,802)	\$3,781 (\$63)	\$5,263 (\$513)	\$5,132 \$7	\$9,613 (\$113)	\$9,700 \$41	\$402 \$0	\$4,946 (\$528)	\$12,974 (\$959)	\$13,223 (\$1,616)
<b>DIF Score Range No. 4 Difference</b>	\$2,599 (\$8)	\$17,140 (\$1,709)	\$4,095 (\$79)	\$5,334 (\$123)	\$5,249 (\$175)	\$10,160 (\$127)	\$18,238 (\$217)	\$7,393 \$0	\$6,307 (\$3,104)	\$10,490 \$150	\$6,927 \$132
<b>DIF Score Range No. 5 Difference</b>	<b>\$3,143</b> (\$95)	\$13,789 (\$87)	\$4,274 (\$46)	\$5,861 (\$177)	\$5,355 (\$130)	\$9,951 (\$25)	\$18,922 (\$2,784)	\$1,004 \$0	\$5,111 \$74	\$16,115 (\$47)	\$17,672 (\$847)
<b>DIF Score Range No. 6 Difference</b>	\$2,460 (\$1)	\$17,386 (\$1,039)	\$4,463 \$17	\$5,556 (\$335)	\$5,942 (\$204)	\$9,333 \$25	\$15,988 (\$2,130)	\$3,637 \$0	\$5,526 \$208	\$14,304 \$290	\$21,892 (\$33,639)
<b>DIF Score Range No. 7 Difference</b>	\$2,999 (\$126)	\$19,476 (\$1,155)	\$5,215 (\$353)	\$5,257 (\$635)	\$5,345 (\$64)	\$9,800 (\$77)	\$17,434 (\$88)	\$68 \$0	\$8,585 \$196	\$20,416 (\$2,837)	\$17,169 (\$3,499)
<b>DIF Score Range No. 8 Difference</b>	\$2,118 \$0	\$19,113 (\$1,714)	\$5,565 (\$58)	\$6,009 (\$54)	\$5,723 (\$991)	\$10,178 \$82	\$13,425 \$0	\$5,613 (\$13,771)	\$5,905 (\$72)	\$15,953 (\$4,453)	\$11,210 \$302
<b>DIF Score Range No. 9 Difference</b>	\$2,661 \$7	\$18,342 (\$3,137)	\$5,101 (\$15)	\$6,536 (\$342)	\$6,103 (\$301)	\$8,430 (\$552)	\$19,645 (\$2,326)	\$436 \$0	\$3,846 \$558	\$13,972 (\$1,966)	\$9,881 \$172
<b>DIF Score Range No. 10 Difference</b>	\$2,419 (\$61)	\$17,680 (\$1,430)	\$5,475 (\$135)	\$6,480 (\$121)	\$5,608 (\$123)	\$9,968 (\$79)	\$18,721 (\$1,017)	\$538 \$0	\$7,513 (\$3,032)	<b>\$21,097</b> (\$5,600)	\$6,849 \$0
<b>DIF Score Range No. 11 Difference</b>	\$2,401 (\$100)	\$20,020 (\$328)	\$6,120 (\$9)	\$7,024 (\$175)	\$5,215 (\$3)	<b>\$11,966</b> (\$103)	\$16,590 (\$5,190)	\$2,272 \$0	\$5,850 (\$3,093)	\$16,906 (\$4,575)	\$27,055 (\$30,570)
<b>DIF Score Range No. 12 Difference</b>	\$2,872 \$8	\$17,428 (\$2,400)	\$5,687 (\$112)	\$6,991 (\$205)	\$6,357 (\$99)	\$10,926 (\$63)	<b>\$22,041</b> (\$3,961)	\$135 \$0	\$6,359 (\$1,825)	\$19,340 (\$1,376)	\$30,673 \$357
<b>DIF Score Range No. 13 Difference</b>	\$2,297 \$6	\$19,727 (\$3,469)	\$6,417 (\$282)	\$7,433 (\$203)	\$6,406 (\$28)	\$10,919 (\$146)	\$16,658 (\$4,186)	\$5,273 \$1,230	\$3,059 \$533	\$14,242 (\$1,041)	\$18,815 (\$6,949)
<b>DIF Score Range No. 14 Difference</b>	\$2,457 (\$148)	\$16,392 (\$359)	\$6,096 (\$643)	\$7,183 (\$75)	\$6,593 (\$339)	\$10,236 (\$17)	\$20,885 (\$1,345)	\$761 \$0	\$6,777 (\$1,496)	\$16,218 (\$1,852)	\$23,354 \$509
<b>DIF Score Range No. 15 Difference</b>	\$2,940 \$4	\$19,534 (\$1,448)	<b>\$7,097</b> (\$329)	\$8,306 (\$309)	\$7,039 (\$429)	\$11,172 (\$73)	\$20,442 (\$5,569)	\$710 \$0	\$4,629 (\$589)	\$17,042 (\$5,475)	\$7,411 \$41,863
<b>DIF Score Range No. 16 Difference</b>	\$2,694 (\$84)	\$12,940 \$30	\$6,545 (\$221)	\$8,217 (\$851)	\$6,978 (\$282)	\$11,693 (\$510)	\$17,256 (\$1,804)	\$5,291 \$0	\$8,122 (\$3,555)	\$18,967 \$678	\$10,659 \$17,036
<b>DIF Score Range No. 17 Difference</b>	\$2,996 (\$18)	\$21,083 (\$8,653)	\$6,576 (\$689)	\$8,191 (\$1,227)	\$6,806 (\$501)	\$11,549 (\$533)	\$18,366 (\$292)	\$425 \$0	\$6,151 \$219	\$13,893 (\$1,000)	\$16,807 \$558
<b>DIF Score Range No. 18 Difference</b>	\$2,135 \$4	<b>\$27,219</b> (\$8,994)	\$6,891 (\$312)	\$9,402 (\$1,140)	\$7,441 (\$294)	\$11,350 (\$493)	\$16,670 (\$5,876)	(\$778) \$1,585	\$6,697 (\$1,409)	\$17,106 (\$5,230)	\$9,951 (\$950)
<b>DIF Score Range No. 19 Difference</b>	\$2,855 (\$109)	\$20,115 (\$4,528)	\$6,121 (\$170)	\$9,915 (\$103)	\$7,926 (\$820)	\$11,859 (\$444)	\$21,471 (\$1,706)	\$7,437 (\$4,861)	\$6,645 (\$879)	\$11,653 \$137	\$33,901 (\$28,389)
<b>DIF Score Range No. 20 Difference</b>	\$2,933 \$0	\$18,929 (\$6,742)	\$6,814 (\$386)	<b>\$10,628</b> (\$523)	<b>\$8,501</b> (\$876)	\$10,669 (\$102)	\$19,547 (\$137)	\$2,307 \$0	<b>\$10,680</b> \$1,999	\$18,376 (\$3,265)	<b>\$34,673</b> (\$4,582)
<b>Did Range No. 20 Result in Higher Average Net Tax Assessment?</b>	No	No	No	Yes	Yes	No	No	No	Yes	No	Yes

Source: TIGTA analysis of RAAS-provided Calendar Year 2015 DIF-selected primary return examinations closed (through Dec. 2017).

<sup>1</sup> Censoring methodology in this context means that the examination results for the top 2 percent of recommended net tax assessments is set to equal to the 98<sup>th</sup> percentile and the bottom 2 percent is set to equal to the 2<sup>nd</sup> percentile.

<sup>2</sup> The difference amounts shown in Figure 2 represent the amount difference between Appendix V, Figures 1 and 2.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

For each examination class, we segregated the data into 20 ranges based on return volume and DIF score. Each range contained about the same number of returns. We computed the average net tax assessment for each range for comparison. Range No. 1 contains returns with the lowest DIF scores among examined returns, and Range No. 20 contains returns with the highest DIF scores. Essentially, we performed similar analysis as shown in Appendix V, Figure 1, but instead of using actual, we used RAAS “censored” examination result amounts. Examination classes with a wide spread of dollar amounts, *e.g.*, high abatements or high adjustments, are more likely to show a difference between Appendix V, Figures 1 and 2.

As Figure 2 shows, even after censoring outliers, more often than not, examined returns with the highest DIF scores still did not result in a higher average net tax assessment than examined returns with lower DIF scores.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Appendix VI**

*Glossary of Terms*

<b>Term</b>	<b>Definition</b>
<b>Audit Information Management System</b>	The AIMS is a computer system used by the SB/SE Division Examination Operations function and others to control returns, input assessments/adjustments to the Master File, and provide management reports.
<b>Calendar Year</b>	The 12-consecutive-month period ending on December 31.
<b>Compliance Data Environment</b>	The Compliance Data Environment project will provide electronic data to other examination systems, such as the Examination Desktop Support System.
<b>Correspondence Examination Automation Support System</b>	The Correspondence Examination Automation Support System is a suite of web-based applications developed to enhance the examination process. Its applications satisfy requests to store/retrieve exam cases to/from a centralized database. It also enables case assignment and transfer between examination groups and batch groups.
<b>Data Center Warehouse</b>	An online database maintained by TIGTA. The Data Center Warehouse pulls data from IRS system resources, such as IRS Collection and Examination files, for TIGTA access.
<b>Direct Examination Staff Year</b>	A DESY is generally equal to 2,000 direct examination hours. Staff years are the total number of revenue agent or tax compliance officer hours divided by fiscal year total hours. Fiscal year total hours are 2,000 (FY 2014), 1,992 (FY 2015), and 2,040 (FY 2016).
<b>Discriminant Function</b>	A mathematical technique used to computer score income tax returns as to examination potential. Examination potential is indicated by a numeric score which is assigned to each return by examination class; the greater the score, the greater the examination potential within each examination class.
<b>Examination Class</b>	A method used to categorize returns by the amount of income or assets.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

Term	Definition
<b>Fiscal Year</b>	Any yearly accounting period, regardless of its relationship to a calendar year. The Federal Government's fiscal year begins on October 1 and ends on September 30.
<b>Goal</b>	The end towards which effort is directed.
<b>Integrated Data Retrieval System</b>	IRS computer system capable of retrieving or updating stored information. It works in conjunction with a taxpayer's account records.
<b>Internal Revenue Manual</b>	The primary, official source of instructions to staff related to the organization, administration, and operation of the IRS.
<b>Metric</b>	A standard of measurement.
<b>Net Tax Gap</b>	The net Tax Gap is the gross Tax Gap less tax that will be subsequently collected, either paid voluntarily or as the result of IRS administrative and enforcement activities.
<b>Revenue Agent</b>	Employees in the Examination function who conduct face-to-face examinations of more complex tax returns such as businesses, partnerships, corporations, and specialty taxes.
<b>Target</b>	A goal to be achieved.
<b>Tax Compliance Officer</b>	Employees in the Examination function who primarily conduct examinations of individual taxpayers through interviews at IRS field offices.
<b>Tax Gap</b>	The estimated difference between the amount of tax that taxpayers should pay and the amount that is paid voluntarily and on time.
<b>Tax Year</b>	A 12-month accounting period for keeping records on income and expenses used as the basis for calculating the annual taxes due. For most individual taxpayers, the tax year is synonymous with the calendar year.



*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

**Appendix VII**

*Management's Response to the Draft Report*



COMMISSIONER  
SMALL BUSINESS/SELF-EMPLOYED DIVISION

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

March 7, 2019

MEMORANDUM FOR MICHAEL E. MCKENNEY  
DEPUTY INSPECTOR GENERAL FOR AUDIT

FROM: *for* Mary Beth Murphy *Tamara Kappard*  
Commissioner, Small Business/Self-Employed Division

SUBJECT: Draft Audit Report – Improvements in Return Scoring and  
Resource Allocation at the Strategic Level Could Enhance  
Examination Productivity (TIGTA #201730031)

Thank you for the opportunity to review and comment on the above subject draft audit report. Annually, the IRS creates an exam plan based on resources requested in the Congressional budget. The exam plan is critical to achieving our overall compliance and coverage objectives. We use a variety of sources to select returns for audit to meet the exam plan. Within our Small Business/Self-Employed division, we group similar sources and types of examinations and call these strategies.

We diligently monitor our progress towards achieving the exam plan throughout the year. As noted in your report, we document trends and deviations from the plan and elevate these observations to management. However, in many instances, we are limited in our ability to control the delivery of returns within certain strategies and take actions to address deviations from the exam plan. Several of our sources of work are taxpayer driven meaning that taxpayers' actions determine how much or how little work we will perform. For instance, taxpayers may amend previously filed returns to reflect claims for refund. We cannot control or limit the number or type of claims for refund filed in a given year. Additionally, events can occur throughout that year that cause us to change our plans. When a natural disaster such as a hurricane occurs, planned work may need to be replaced with work that may not have been in the original workplan. This may shift work from one strategy to another. As we recognize shifts between strategies, whether intentional or not, we document them and take corrective action where possible and appropriate.

We appreciate your recognition of our improvements to survey and no change rates. Specifically, we took steps to streamline the inventory ordering process and introduced tools to prioritize delivering those returns with the highest audit potential. These actions enhanced our ability to utilize our exam resources efficiently and reduced taxpayer burden by more effectively selecting the taxpayers with the highest potential for non-



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

2

compliance. Based on these process improvements, surveys decreased from 26% to 15% and our no change rate decreased from 10% to 8% during Fiscal Year (FY) 2014 to FY 2016.

We do not agree with your analysis or findings regarding the discriminant function (DIF) strategy. DIF uses predictive analytical technique models to score returns at the time of filing. Generally, returns with higher scores have a greater likelihood of having significantly understated tax. While no return selection method correctly identifies non-compliance in every case, our analysis over the years repeatedly indicates that DIF successfully alerts us to returns with a high potential on non-compliance. DIF models are developed, evaluated and used operationally to focus our examination resources on only the highest scored returns. To properly assess DIF's performance, the results of highly scored returns should be compared to the results of lower scored returns and TIGTA's analysis did not make this assessment.

The National Research Program (NRP), like DIF, is a strategy focused on the entire return population. Approximately 14,000 NRP returns selected each year from Tax Year 2006 to Tax Year 2015 were selected using a rigorous, statistical, stratified sampling approach. This is done so that, in addition to establishing the correct amount of tax for a given return, the information gathered can be used for other purposes that require data reflective of compliance of the entire population of the more than 150 million returns that are filed each year. In addition to informing and improving DIF models, NRP results are used to estimate the tax gap and estimate improper payments.

We do not agree with the potential increased revenue of \$262.5 million. This methodology assumes that performance will be the same for all returns within each strategy and does not reflect other Exam goals of compliance and coverage. The outcome measure also assumes that all work within the workplan will be completed as planned. As previously discussed, uncontrollable deviations may occur.

Attached is a detailed response to address your recommendations. If you have any questions, please contact me or Brenda Dial, Director, Examination Operations, SB/SE Division.

Attachment



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

Attachment

**RECOMMENDATION 1:**

The Commissioner, Small Business/Self-Employed Division, should monitor performance goals and take corrective actions as necessary to achieve targets specific to staffing resources, e.g., revenue agents and tax compliance officers, and establish actionable performance goals with a range for the number of closures for each strategic priority to ensure that they provide balanced coverage.

**CORRECTIVE ACTION:**

We do not agree with this recommendation. Similar workstreams are grouped together by strategy. Based on estimates of inventory volumes and delivery timeframes, the approximate number of closures and hours are generated for the cases in each strategy. Examination does not control the delivery of returns within many strategies, such as NRP, Refund Claims, and Special Enforcement Program & Fraud. Taxpayer behavior drives the number of refund claims received and the IRS can neither control the number nor types of claims filed. As noted in your report, we already document trends and deviations from the plan and elevate these observations to management. We will continue to monitor and report closures by strategy and return category in an effort to identify causes of any shortfalls or surplus closures and take action where possible and appropriate.

**IMPLEMENTATION DATE:**

N/A

**RESPONSIBLE OFFICIAL:**

N/A

**CORRECTIVE ACTION MONITORING PLAN:**

N/A

**RECOMMENDATION 2:**

The Chief Research and Analytics Officer should evaluate the examination results to determine whether refinements could improve the correlation of the DIF scores and net tax assessments for individual returns examined by revenue agents and tax compliance officers.

**CORRECTIVE ACTION:**

We disagree with this recommendation. We regularly update the DIF scoring models, striving to improve upon the already effective models. Based upon our routine reviews of the DIF scoring models and independent of this audit, 10 of the 11 examination classes that were reviewed in the report have already been updated and implemented since the processing year that was



---

*Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity*

---

2

evaluated. Therefore, the report's specific results do not provide any actionable information relevant to the current DIF models. Although we remain committed to continuing our efforts to evaluate DIF performance and improve the performance of the DIF scoring models, we do not agree that the analysis in this report supports any additional corrective action.

DIF is highly effective at identifying returns in the population that are likely to have a significant change. Analysis of DIF selected returns and examinations of a representative sample of returns selected as part of NRP establish that the average tax change for examined returns with high DIF scores is significantly higher than the expected tax change for examined returns with lower DIF scores. The evaluation methods in this report, in some instances relying upon a single return, lead to conclusions that are not generalizable and therefore we cannot rely upon those methods or conclusions for evaluating predictive models. The fact that the return with the highest tax change may not necessarily have the highest DIF score reflects the inherent uncertainty with any statistical or probabilistic modeling.

**IMPLEMENTATION DATE:**

N/A

**RESPONSIBLE OFFICIAL:**

N/A

**CORRECTIVE ACTION MONITORING P**

N/A