TREASURY INSPECTOR GENERAL FOR TAX ADMINISTRATION

# Individual Returns With Large Business Losses and No Income Pose Significant Compliance Risk

September 28, 2020

Reference Number: 2020-30-056

TIGTACommunications@tigta.treas.gov | www.treasury.gov/tigta | 202-622-6500

# HIGHLIGHTS: Individual Returns With Large Business Losses and No Income Pose Significant Compliance Risk

**Final Audit Report issued on September 28, 2020**
**Reference Number 2020-30-056**

## Why TIGTA Did This Audit

For Tax Years 2011 through 2013, the IRS estimated the annual gross Tax Gap to be $441 billion, with underreporting contributing approximately $352 billion. Of this underreporting noncompliance, $245 billion is attributed to individual income tax underreporting, a large portion of which ($110 billion, or 45 percent) is attributed to individual business income reported on individual income tax returns.

This audit was initiated to determine whether the IRS's systemic and nonsystemic methods and processes for scoring, selecting, and delivering tax returns of nonfarm sole proprietors with Form 1040, *U.S. Individual Income Tax Return*, Schedule C, *Profit or Loss From Business (Sole Proprietorship)*, are effective at identifying potential noncompliance.

## Impact on Taxpayers

According to recent statistics, less than 1 percent of returns filed during Calendar Year 2017 were examined in Fiscal Year 2018. Given its limited resources, it is essential for the IRS to have a workload selection methodology in place that effectively identifies tax returns that present the largest risks to tax compliance for examination.

## What TIGTA Found

For Tax Years 2013 through 2017, a total of 29,378 field examinations were conducted and closed on Forms 1040 that have at least one Schedule C with no gross receipts and no profit. Of those examinations, 1,142 returns contain either a single Schedule C loss that was equal to or greater than $100,000 or multiple Schedule C forms that showed combined losses that were equal to or greater than $100,000. The 1,142 returns had an average examination assessment of $53,183, which was greater than the examination results of seven of the 10 Small Business/Self-Employed Division's (SB/SE) Field Examination function strategies. Additionally, the results of SB/SE's Compliance Initiative Projects have demonstrated that these types of returns are more productive than Discriminant Function (DIF)-selected returns. Yet, the SB/SE Division utilizes DIF to select most of its returns for examination. Additional research is needed to determine if the IRS could focus its resources on more productive examinations.

Finally, the last update to internal controls and procedural documents related to the DIF Development Process was in May 2014. The DIF scoring models should be updated more frequently to ensure that the IRS continues to have the best tools available to select the most productive returns to examine.

## What TIGTA Recommended

TIGTA recommended that the Commissioner, SB/SE Division, conduct further research on returns with Schedule C attached with no gross receipts and more than $100,000 in losses to evaluate these returns as a new workstream and collaborate with the Research, Applied Analytics, and Statistics (RAAS) organization to improve any issue-based workload selection methodologies to enhance the return selection process and improve operational outcomes. In addition, RAAS should establish, document, and implement procedures to ensure that documentation associated with internal controls and procedures around the DIF scoring methodology and model development process are reviewed and updated regularly and consistently.

The IRS disagreed with two of the recommendations and agreed with one recommendation. The IRS did not agree to conduct additional research on returns with at least one Schedule C attached, no gross receipts, and more than $100,000 in losses. The IRS agreed to develop a document that outlines the program process in a sequential timeline format to add clarity to the processes as well as adequately document efforts to evaluate potential DIF alternatives.

TREASURY INSPECTOR GENERAL
FOR TAX ADMINISTRATION

**U.S. DEPARTMENT OF THE TREASURY**

**WASHINGTON, D.C. 20220**

September 28, 2020

**MEMORANDUM FOR:** COMMISSIONER OF INTERNAL REVENUE

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

**SUBJECT:** Final Audit Report – Individual Returns With Large Business Losses and No Income Pose Significant Compliance Risk (Audit # 201830016)

This report presents the results of our review to determine whether the Internal Revenue Service's systemic and nonsystemic methods and processes for scoring, selecting, and delivering tax returns of nonfarm sole proprietors with Form 1040, *U.S. Individual Income Tax Return*, Schedule C, *Profit or Loss From Business (Sole Proprietorship)*, are effective at identifying potential noncompliance. This review is part of our Fiscal Year 2020 Annual Audit Plan and addresses the major management and performance challenge of *Improving Tax Reporting and Payment Compliance*.

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

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).

# Table of Contents



## Appendices

# Background

A sole proprietor actively involved in furthering the interests of a business whose primary purpose is to generate revenue/profit must file a Form 1040, *U.S. Individual Income Tax Return*, Schedule C, *Profit or Loss From Business (Sole Proprietorship)*.[1] A separate Schedule C must be filed for each business owned.

An individual's Form 1040 with a Schedule C attached can present challenges to the Internal Revenue Service's (IRS) efforts to verify tax compliance due to the lack of third-party data associated with income and potentially complex business expense deductions. IRS research has demonstrated that there is a visible link between reporting compliance and third-party information reporting. When there is no tax withholding or little or no information reporting, the IRS estimates net misreporting to be 55 percent.

An individual sole proprietor's receipts and expenses are used to calculate business profits or losses, which are then combined with income, deductions, and credits from other sources that are reported elsewhere on the Form 1040 to compute a taxpayer's overall individual tax liability. Identifying which of a sole proprietor's payments qualify as business expenses and can be deducted, or whether an appropriate amount of receipts and profits were reported, can be complex, and the noncompliance has the potential to be significant, both individually and for this portion of the taxpayer population as a whole.

Specifically, for Tax Years (TY) 2011 through 2013, the IRS estimated the annual gross Tax Gap to be $441 billion, with underreporting contributing approximately $352 billion.[2] Of this underreporting noncompliance, the majority, or $245 billion, is attributed to individual income tax underreporting, a large portion of which ($110 billion, or 45 percent) is attributed to individual business income reported on individual income tax returns. Similarly, the estimated annual Tax Gap for TYs 2008 through 2010 showed that $125 billion of the $264 billion (47 percent) attributed to individual underreporting was associated with individual business income.[3]

## Impact of examinations on noncompliance

Courts have also found that a consistent, sizeable underreporting of income is also one of the "badges of fraud" by which the courts seek to assess whether a taxpayer has committed tax fraud.[4] To the extent that the IRS can systemically identify consistent, sizeable underreporting, it would assist the IRS in addressing the most significant noncompliance. By selecting and treating individual tax returns that have a high risk of noncompliance, IRS examinations could potentially bring taxpayers back into compliance. Past IRS research also suggests that the indirect effect of examinations on overall tax compliance is significantly more than the impact of an individual examination, *i.e.*, the impact of a single examination has ripple effects on tax

[1] The Schedule C is a tax form used to report annual income or loss from a sole proprietorship or single-member Limited Liability Corporation. This form is submitted as an attachment to the Form 1040.
[2] IRS, Publication 1415 (Sept. 2019). A tax year is the 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.
[3] IRS, Publication 1415, *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008–2010* (May 2016).
[4] *McGraw v. Comm'r*, 384 F.3d 965, (8th Cir. 2004).

compliance that extend beyond just that examination. Specifically, the deterrent effect of audits started in TY 1991 was more than 11 times as large as the average adjustment directly proposed by audits closed that year.[5]

IRS Policy Statement 4-21 states that the primary objective in selecting returns for examination is to promote the highest degree of voluntary compliance on the part of taxpayers. This policy requires the IRS to exercise professional judgment to select sufficient returns of all classes of returns to assure all taxpayers equitable consideration while using available experience and statistics to make the most efficient use of examination staff and other resources in an effort to decrease the Tax Gap and bring taxpayers into compliance.

The IRS employs multiple strategies to deliver its examination inventory across its various business units. The Small Business/Self-Employed (SB/SE) Division's Examination function, which processes and evaluates the examination potential for the majority of the tax returns with Schedule C, used 10 strategic priorities to group similar examination work between Fiscal Years (FY) 2014 and 2018.[6] Each strategy provides a different type of inventory. Figure 1 shows the total number of examinations of individual tax returns conducted and closed by the SB/SE Division Field Examination function for each strategy during FYs 2014 through 2018.

[5] IRS Publication 1916, *The Determinants of Individual Income Tax Compliance: Estimating The Impacts of Tax Policy, Enforcement, and IRS Responsiveness* (Nov. 1996). This analysis was conducted on individual taxpayer data over a 10-year period, 1982 through 1991, aggregated to the State level.

[6] The workstreams are listed on the SB/SE Division U.S. Monitoring and Performance Reports. In FY 2018, the High-Income and High Wealth Taxpayers strategy was revised and the name of the strategy was changed to Nonfilers. Additionally, in FY 2018, a new strategy, Claims, was added; however, because it was only reported one year (FY 2018), we did not include it in the chart in Figure 1. There were 11,385 total closures under the Claims strategy in FY 2018. A fiscal year is 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.

**Figure 1: Summary of Examination Closures of Individual Returns by Strategy**

| Strategy | FY 2014 | FY 2015 | FY 2016 | FY 2017 | FY 2018 | Total Closures | Percentage |
|---|---|---|---|---|---|---|---|
| Discriminant Function (DIF) Scoring Methodology | 119,842 | 111,821 | 87,799 | 72,654 | 64,328 | 456,444 | 45% |
| Other Priority Programs | 25,459 | 24,858 | 25,235 | 30,292 | 12,207 | 118,051 | 12% |
| Promoters and Return Preparers | 37,061 | 24,439 | 20,063 | 14,321 | 10,454 | 106,338 | 11% |
| National Research Project | 14,222 | 14,816 | 23,144 | 18,991 | 15,553 | 86,726 | 9% |
| Offshore | 16,985 | 17,391 | 17,065 | 15,373 | 9,450 | 76,264 | 8% |
| All Other | 20,702 | 14,630 | 12,454 | 9,828 | 10,555 | 68,169 | 7% |
| Unreported and Underreported Income | 5,265 | 9,938 | 11,605 | 7,620 | 2,754 | 37,182 | 4% |
| High-Income and High Wealth Taxpayers/Nonfilers | 7,875 | 8,250 | 9,910 | 6,364 | 4,071 | 36,470 | 4% |
| Abusive Transactions and Technical Issues | 4,983 | 3,882 | 2,976 | 1,757 | 1,842 | 15,440 | 2% |
| Special Enforcement Program and Fraud | 1,995 | 1,885 | 1,712 | 1,685 | 1,299 | 8,576 | 1% |
| **Total** | **254,389** | **231,910** | **211,963** | **178,885** | **132,513** | **1,009,660** | **100%[7]** |

*Source: Treasury Inspector General for Tax Administration (TIGTA) summary of SB/SE Division Examination U.S. Monitoring and Performance Reports for FY 2014 through FY 2018.*

## DIF Scoring Methodology

While several of the SB/SE Division's strategies may identify audit risk in a Schedule C and select the return for examination, the primary strategy used by the SB/SE Division's Field Examination function for Schedule C returns is the DIF Scoring Methodology (hereafter referred to as "DIF"). DIF's tax return scoring models are a formula that automatically measure and grade the risk of noncompliance and subsequently assign a score to the return based on the potential for overall tax change to the taxpayer's return.[8]

The SB/SE Division Field Examination function identifies its discretionary inventory from tax returns based on descending DIF score, with high scores having the most overall potential for tax change. Although DIF scoring models incorporate an objective analysis of certain line items or variables, or a combination thereof, as noted earlier, the objective is to assign a score to the return based on the potential for overall tax change and not to flag potentially high-risk line item issues. The Examination function's downstream processes include manual review of the tax returns in order to identify line items or issues for examination or to potentially accept the return as filed.

Prior to DIF scoring, all individual returns were reviewed for audit potential through a manual review. The IRS has described this process as having hundreds of the most experienced examiners sort through millions of returns each year by hand. They selected returns using general guidelines and according to their judgment and professional experience. However, this practice resulted in taking the most experienced examiners away from conducting examinations

[7] Amounts may not total due to rounding.

[8] DIF was used initially to score individual tax returns only, but it is currently used to score Forms 1120, *U.S. Corporation Income Tax Return;* 1120S, *U.S. Income Tax Return for an S Corporation;* and 1065, *U.S. Return of Partnership Income,* as well as all Forms 1040.

of returns as well as a lack of uniformity in the selection of returns for examination. Due to its resource-intensive nature, not all returns could be screened during this process, which caused inequitable treatment in the selection process.

In 1963, a computer system was developed to identify tax returns for examination with criteria developed using examiner experience. However, the system identified more returns for examination than the IRS's resources could accommodate, and noncompliance could not be readily identified for the groups of returns selected. Through the remainder of the 1960s, the IRS continued to explore methods to identify tax returns for examination and implemented DIF in 1971.

DIF formulas are developed from results of examinations of a sample of tax returns statistically selected to be representative of the population of filed tax returns.[9] Once the models are developed, they are then implemented so that returns can be scored during the processing of a return. As returns are received and processed, a DIF score is calculated and assigned to every return. In theory, the higher a tax return's DIF score, the greater the probability of a tax change if the return is examined. Those returns with the highest DIF scores are made available to the Field Examination function for further classification and assignment, as needed. However, a recent TIGTA report noted that, more often than not, examined returns with the highest DIF scores did not result in the highest average net tax assessment when compared to examined returns with lower DIF scores.[10]

DIF remains the SB/SE Division's primary method to identify its discretionary field examination inventory. The IRS uses the following three measures to develop DIF scoring for each of the Field Examination function's activity codes:

- The no-change rate. The no-change rate is the proportion of examined returns with little or no change in tax liability.
- The average tax change. The average tax change is the average amount of recommended tax adjustment of examined returns.
- The "hit rate." The hit rate is the number of returns with a tax change above a certain threshold to the number of returns examined.[11]

When developing and evaluating potential models, the evaluation compares potential models to the existing model on these three measures and looks for improvements in the measures. A limited number of staff within the IRS's Research, Applied Analytics, and Statistics (RAAS) organization Compliance Modeling Lab review the scoring of a sample of returns to ensure that return scoring is working properly.[12]

While tax returns selected for field examination based on their DIF score is the largest single source of returns selected for examination in the field, total field examination inventory is derived from multiple strategies. The allocation of inventory for each strategy changes each fiscal year according to the Examination function's workplan, which is dependent on resources

[9] Currently this sample of returns is obtained through the IRS's National Research Project.

[10] TIGTA, Ref. No. 2019-30-024, *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity* (Apr. 2019). The IRS disagreed with TIGTA's analysis and findings regarding the DIF.

[11] Activity code identifies the type and condition of the tax return selected for audit.

[12] A processing year is the calendar year in which the tax return or document is processed by the IRS.

available for the fiscal year. DIF and All Other strategies are built into the examination plan after allocations have been made to the other strategic priorities.

While the number of closures for returns in the DIF strategy has been decreasing from a high of 119,842 closures in FY 2014 to 64,328 closures in FY 2018, it continues to comprise the single greatest number of closures compared to all the alternative strategies. As shown previously in Figure 1, 45 percent of all examinations of individual returns conducted by the SB/SE Division Field Examination function were from the DIF strategy.

# Results of Review

## Returns With at Least One Schedule C Reporting No Income and Substantial Schedule C Losses Pose Significant Compliance Risk

Our review of closed SB/SE Division field examinations found that the average assessment amount of tax returns with at least one Schedule C reporting only business losses was greater than the average assessment amounts of the majority of the SB/SE Division Field Examination function strategies. Specifically, we identified 29,378 TYs 2013 through 2017 Forms 1040 having at least one Schedule C with no gross receipts and no profit.[13]

- 1,142 returns contained a loss on at least one Schedule C that was equal to or greater than $100,000.
- 28,236 returns contained a loss on at least one Schedule C that was less than $100,000.

Our analysis of the 1,142 returns that show a loss equal to or greater than $100,000 found that the average assessment per examination was $53,183, which is greater than the average assessment for seven of the 10 SB/SE Division Field Examination function strategies. For the remaining three strategies that have average assessment amounts greater than $53,183, there appear to be potential outlier amounts for one of the strategies in FYs 2017 and 2018 making the average assessment amount for this strategy significantly higher.[14] For the 28,236 returns that show a loss of less than $100,000, the average assessment per examination was $9,422. This average assessment amount per examination is greater than two of the 10 strategies. Figure 2 shows how the average assessment amounts from our data analysis compare to those of the strategies.

[13] If a taxpayer filed multiple Schedules C, there may be a profit showing on one of the multiple Schedules C attached to the return; however, for our criteria, at least one of the multiple Schedules C showed no gross receipts and was reporting a loss. Additionally, these TYs 2013 through 2017 individual tax returns were limited to the tax returns processed in Processing Years 2014 through 2018, as of February 2019, and examined through February 2019.

[14] For the Abusive Transactions and Technical Issues strategy, the average assessment amounts for FYs 2014 through 2018 were $25,448, $32,613, $50,076, $152,203, and $163,828, respectively. The amounts increase each fiscal year; however, the amounts for FYs 2017 and 2018 appear to be significantly higher. For example, there was a 204 percent increase from FY 2016 to FY 2017 and a 227 percent increase from FY 2016 to FY 2018.

**Figure 2: Comparison of the Average Assessment Amounts**

| IRS Strategy and TIGTA Analysis of Schedule C | Average Assessment Amount |
|---|---|
| Abusive Transactions and Technical Issues | $84,854 |
| Special Enforcement Program and Fraud | $70,288 |
| High-Income and High Wealth Taxpayers/Nonfilers | $53,929 |
| Schedule C Loss of $100,000 or More | $53,183 |
| All Other | $27,788 |
| Other Priority Programs | $27,385 |
| Unreported and Underreported Income | $25,480 |
| Offshore | $17,635 |
| DIF | $15,185 |
| Schedule C Loss of Less Than $100,000 | $9,422 |
| Promoters and Return Preparer | $6,800 |
| National Research Project | $5,952 |

*Source: Strategy average assessment amounts are based on FYs 2014 through 2018 U.S. Monitoring and Performance Reports. Schedule C Loss average assessment amounts are based on TIGTA's analysis of tax account information from the Individual Return Transaction File database for Processing Years 2014 through 2018 matched against the Audit Information Management System database.*

In addition, we identified 33,176 individual returns with at least one Schedule C reporting loss equal to or greater than $100,000 that were never considered or selected for examination (referred to as non-selected returns).[15] As noted previously, individual tax returns with at least one Schedule C reporting only losses show a high compliance risk and a high return on investment. As such, the IRS should focus on these types of cases to address taxpayers with the highest compliance risk and increase the efficient use of its limited resources. Figure 3 provides a stratification of the 33,176 non-selected returns based on the loss amounts claimed on the Schedules C.

[15] These 33,176 cases were identified as of February 2019. The IRS provided an analysis on these cases after we completed our fieldwork. We did not validate the IRS's analysis; however, the IRS indicated that some of these cases may have since been examined, are not SB/SE taxpayers, or have collectability indicators.

**Figure 3: Non-Selected Individual Returns Stratified by Loss Amounts Reported on Schedules C**

| Range of Loss Amount | Number of Tax Returns |
|---|---|
| $100,000 to $249,999 | 24,095 |
| $250,000 to $499,999 | 5,646 |
| $500,000 to $749,999 | 1,573 |
| $750,000 to $999,999 | 625 |
| $1,000,000 to $10,000,000 | 1,163 |
| Greater Than $10,000,000 | 74 |
| **Total** | **33,176** |

*Source: TIGTA analysis of tax account information from the Individual Return Transaction File and Audit Information Management System databases.*

However, SB/SE Division management indicated that the purpose of its strategic priorities is not based solely on whether they are productive in terms of dollars assessed. It is logical that the amount of dollars assessed is not the only characteristic considered in identifying productive workload. However, we believe that these returns are high risk and can be easily identified.

## IRS Compliance Initiative Project (CIP) results also show greater average yield for the Schedule C population

Since May 2011, the IRS has conducted three CIPs related to tax returns with characteristics similar to those we previously analyzed. The CIPs covered three different SB/SE Division areas: North Atlantic Area, Southwest Area, and Western Area. The criteria used to select the tax returns for each CIP differed slightly among each area; however, all of the CIPs focused on Form 1040 tax returns having at least one Schedule C. The overall strategy addressed by all three of these CIPs was to increase compliance among SB/SE Division taxpayers. The performance measures of the CIPs were based on dollars adjusted per hour (also referred to as yield), dollars adjusted per return, and the no-change rate. All three of the CIPs showed favorable results when compared to DIF-related audit results during the same time frame.

For example, the CIPs resulted in average yields that were greater than the average yields of DIF-related audits. Specifically, the average yields for the North Atlantic, Southwest, and Western Areas were 49 percent, 11 percent, and 55 percent, respectively, greater than DIF. Figure 4 shows the comparison of CIP results to the DIF.

**Figure 4: Comparison of CIP Results to DIF[16]**

| **Area** | Date | Criteria | **CIP** | | | **DIF** | | |
|---|---|---|---|---|---|---|---|---|
| | | | Yield | Assessment Per Return | No-Change Rate | Yield | Assessment Per Return | No-Change Rate |
| **North Atlantic** | May 2011 to January 2015 | *******2*******<br>*******2*******<br>*******2**********<br>*******2**********<br>*******2******* | $857 | $11,294 | 6% | $574 | $12,404 | 10% |
| **Southwest** | December 2012 to December 2015 | *******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2******* | $562 | $18,939 | 6% | $507 | $20,483 | 18% |
| **Western** | July 2014 to June 2016 | *******2*********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2**********<br>*******2******* | $1,001 | $15,323 | 8% | $644 | $13,807 | 9% |

*Source: TIGTA analysis of SB/SE Division CIP reports and SB/SE Division Examination U.S. Monitoring and Performance Reports for FY 2012 through FY 2016.*

The IRS generally describes the results of CIPs as being specific to a certain geographic area, as the SB/SE Division Examination Field function is organized geographically and each area may establish its own CIPs. While the results associated with the IRS's prior work described previously may be dated, the results establish a pattern which indicates that the IRS, specifically the SB/SE Division, has the knowledge and ability to identify more productive returns than what may be found through its strategic priorities and, specifically, its discretionary DIF strategy. During our discussions, SB/SE Division Examination function management indicated that they had no knowledge or evidence for the CIPs becoming permanent workstreams.

[16] All values in Figure 4 were documented in the IRS's CIP reports except the DIF Assessment Per Return amount and the DIF No-Change Rate for the North Atlantic Area and the Western Area. Also, the DIF Yield for the Western Area was not documented in the IRS's CIP reports. We determined these values using the IRS U.S. Monitoring Reports, which shows the results of Field Examination function case closures of individual returns with a Schedule C worked under the DIF strategy. We then calculated the averages based on the results shown in the U.S. Monitoring Reports for each of these categories over a similar time frame of the CIPs. Specifically, DIF values for the North Atlantic Area were based on FYs 2012 through 2015, and DIF values for the Western Area were based on FYs 2014 through 2016.

## The IRS should consider specific compliance gaps in conjunction with DIF when considering potential inventory

The DIF strategy allows the SB/SE Division to provide discretionary inventory to its examiners once all other priorities are assigned; however, the SB/SE Division may not be using its resources to the highest potential given the manner it is utilizing DIF. TIGTA recently reported that, even though DIF identifies returns with examination potential, more often than not, individual returns with the highest DIF scores do not result in a higher net tax assessment for most examination classes.[17] The results of successful or productive CIPs and past compliance results should be used in conjunction with or as a complement to DIF to identify more productive inventory.

We spoke with SB/SE Division Examination function management about their perspective on the DIF and resulting inventory. While IRS management believes that tax returns delivered through DIF strategy are an important component to their inventory, management intends to backfill this discretionary work with other, more productive, cases. When provided with preliminary results of our analysis and perspective associated with the segment of potentially more productive returns, IRS management responded that they are consistently finding new inventory sources that require a reallocation of resources allotted to DIF. IRS management believes that the expansion of other inventory sources/strategies coupled with the reduction in overall staffing have resulted in less DIF strategy work despite the importance of DIF to promote voluntary compliance.

They also described current efforts to implement the Enterprise Planning Scenario Tool and the Return Order and Delivery System to identify better discretionary DIF inventory for assignment.[18] The individual returns with a Schedule C included within our analysis were examined prior to implementation of these tools. As such, we did not evaluate the impact these tools may have on the Examination function's inventory.

While DIF scoring models incorporate analysis and consideration of various line items, or combination of line items, the IRS stressed during this audit that the purpose of DIF is not to identify any one particular type of noncompliance or issue but to score tax returns on the whole based on their potential for tax change. DIF is not a "filter" designed to identify returns that meet a single, specific criteria. Because DIF was developed to score tax returns on the whole, there are limitations in the value that it has to directly identify the specific types of tax return issues we described previously.

The SB/SE Division management believes that issue-based strategies have had little success in the past. They also believe that the SB/SE Division is not configured to transition to an issue-based examination approach.

---

[17] TIGTA, Ref. No. 2019-30-024, *Improvements in Return Scoring and Resource Allocation at the Strategic Level Could Enhance Examination Productivity* (Apr. 2019).

[18] The Enterprise Planning Scenario Tool utilizes the available hours for DIF returns, available return inventory, historic results (hours per return and outcome of examinations), and staffing in an effort to optimize selected objectives such as Dollars Recommended, Dollars Collected, and No-Change Rate. It was first used in FY 2018 for tax compliance officer inventory, and in FY 2019, it was used in some revenue agent and all tax compliance officer inventories. Management expected the tool to be heavily relied on in FY 2020. The Return Order and Delivery System considers inventory and staff that are available in order to allocate inventory. Tax compliance officers primarily conduct examinations of individual taxpayers through interviews at IRS field offices, and revenue agents conduct face-to-face examinations of more complex tax returns such as businesses, partnerships, corporations, and specialty taxes, *e.g.*, excise tax returns.

Two of the IRS's business units (the Large Business and International and the Tax Exempt and Government Entities Divisions) have moved in the direction of becoming more issue based in their audit selection techniques. While there may be barriers to the SB/SE Division implementing a more issue-based examination approach, DIF may be a useful enhancement or complement to assist in prioritizing returns if the SB/SE Division were to identify a subset of taxpayer returns with a particular issue or with a common compliance gap. For example, the IRS may be able to evaluate the risk associated with the issues and line items that are incorporated within DIF's scoring models or consider whether DIF could be updated or retooled to identify particular issues on a segment of taxpayer returns.

In addition, during our discussions with RAAS organization officials regarding the DIF's development process, they noted that there are significant challenges in maintaining the IRS's National Research Project, including recent, significant tax law changes and resource challenges, which could potentially lead to degradation of DIF because DIF formulas are updated based on National Research Project data results.[19] It is imperative that the SB/SE Division Examination function make an effort to assess its current DIF strategy and discretionary workstream to determine whether it may be able to supplement that with more productive work.

The Commissioner, SB/SE Division, should:

**Recommendation 1:** Conduct a national CIP on tax returns with at least one Schedule C attached, no gross receipts, and more than $100,000 in losses to evaluate it as a new strategy or workstream. The IRS should use the methods and results from this CIP to consider other compliance gaps, past CIPs, or issues to ensure that the SB/SE Division is selecting the most productive returns for examination.

**Management's Response:** The IRS disagreed with this recommendation. In their response, IRS management stated that they have significant coverage of Schedule C returns, and their various workload selection methodologies identify the most productive returns for examination while balancing coverage and other noncompliance objectives. In addition, the IRS does not believe there is justification to commit significant additional resources to the population TIGTA identified.

**Office of Audit Comment:** Our audit shows that CIPs have the potential to identify more significant pockets of noncompliance than the reliance on DIF. Our audit results indicate that tax returns with at least one Schedule C attached that report significant losses and no income require IRS attention. Additionally, disregarding our recommendation for further research of the compliance issue identified is inconsistent with SB/SE Division's focus on leveraging research and data analytics to detect emerging issues.

As noted in our report, individual tax returns with at least one Schedule C attached reporting no gross receipts claiming $100,000 or more in losses pose a significant compliance risk. In addition, our review found that the average assessment amount of tax returns with at least one Schedule C reporting only business losses of $100,000 or more was greater than the average assessment

[19] The National Research Project 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.

amounts of the majority of the SB/SE Division Field Examination function strategies.

Recommendation 2: Collaborate with the RAAS organization to improve any issue-based workload selection methodologies developed under Recommendation 1 by considering use of DIF in conjunction with other methods to enhance the return selection process and improve operational outcomes.

**Management's Response:** The IRS disagreed with the recommendation. The IRS stated that this recommendation is not applicable because the IRS does not agree with Recommendation 1.

**Office of Audit Comment:** We believe that the SB/SE Division should collaborate with the RAAS organization to leverage the results of its research and data analysis to improve its workload selection methodologies. As we point out in our Office of Audit Comment to Recommendation 1, these tax returns pose a significant compliance risk. We believe these returns should be evaluated to ensure that the SB/SE Division is selecting the most productive returns for examination.

## Discriminant Function Internal Control Documents and Alternatives to the Discriminant Function Scoring Methodology Should Be Updated More Frequently

Internal control documents associated with the DIF development process have not been updated in several years, and alternatives to the DIF need to be considered to ensure accuracy and productivity. The most recent updates to the DIF Development Process control documents were in May 2014. As such, the current DIF development methodology may not include updated characteristics to assess the current risk environment in tax administration accurately. In addition, it has been approximately 13 years since the IRS has evaluated how other analytical techniques and alternative software compare to discriminant analysis.

### RAAS needs to update DIF internal control documents more frequently

During this review, we received copies of several internal control and procedural documents, including two IRS DIF Development Process documents dated January 2005 and May 2014. While both IRS DIF Development Process documents provide an overview of the history of the DIF's process and details of the steps involved in the development of the DIF's formula and scoring methodology, the one dated May 2014 provides for a more detailed explanation of the steps involved in the development of the DIF formula and scoring methodology. In addition, the document includes discussion on the DIF's strengths and weaknesses.

Although the IRS was able to provide these internal control documents, nearly 10 years had passed between iterations of the IRS DIF Development Process documents we were provided. Any number of events have the potential to affect the development of the DIF's scoring process, including changes to regulations or tax laws, or impact from changes associated with the National Research Project, from which DIF is derived.

The development methodology is part of the IRS's internal control system for DIF scoring. An effective internal control system can help Federal agencies achieve their missions and objectives and improve accountability. The Government Accountability Office's *Standards for Internal Control in the Federal Government*, also known as the Green Book, states that internal controls comprise the plans, methods, and procedures used to meet an entity's mission, goals, and objectives, which support performance-based management. Internal controls help agency program managers achieve desired results and provide reasonable assurance that program objectives are being achieved through, among other things, effective and efficient use of agency resources. In addition, internal control is not one event, but rather an ongoing series of actions and activities that occur throughout an entity's operations.

Only a limited number of staff have access to DIF development details. Given the significant impact that the DIF's development process can have on many taxpayers, as well as the limited number of RAAS organization staff involved in the process, management needs to establish a baseline to monitor and update the internal control system and associated documents regularly and to ensure that policies and procedures are accurate, thorough, and incorporate any necessary changes from year to year. This would help ensure that institutional knowledge is retained as changes in staffing occur in the future.

## Alternatives to the DIF scoring methodology should be evaluated more frequently

Since implementation of DIF's methodology in 1971, there have been several studies of the software's ability to identify productive returns and determine if better alternative scoring models exist. During our fieldwork in March 2019, the RAAS organization approved funding for its most recent evaluation of other analytical techniques and how alternative software compares to discriminant analysis. While this update is a positive development, it still reflects a 13-year time gap since the prior evaluation.

An internal control document dated June 2014 states that these studies have produced mixed or inferior results when compared to DIF. See Appendix II for a summary of research and studies we were provided that were intended to consider alternative scoring techniques. Several of the studies recommended that additional work be performed, but it is not evident from the documentation we were provided whether these follow-on studies or evaluations were ever completed. From the dates associated with these evaluations of DIF, the studies have not been performed regularly.

One of the RAAS organization's core values is innovation, which is described as an effort to "…transform and continuously improve processes, and conduct cutting-edge research to improve tax administration and IRS operations." With more than a decade between evaluations, the IRS has the potential to miss significant changes in how it organizes and prioritizes examination work, and subsequently how alternative methodologies may compare, favorably or unfavorably, to DIF.

According to recent statistics, taxpayers filed nearly 196 million returns during Calendar Year 2017, of which 991,168 (less than 1 percent) returns were examined in FY 2018.[20] The SB/SE Division Field Examination function derived almost one-half of its examination closures,

[20] IRS Publication 55B, *Internal Revenue Service Data Book, 2018*, p. 23 (May 2019). In general, examination activity is associated with returns filed in the previous calendar year; though the IRS has three years from the filing date to conduct the examination.

approximately 49 percent, through its discretionary DIF strategic priority.[21] Given its limited resources, it is essential for the IRS to have in place a workload selection methodology that effectively identifies the most productive tax returns for examination.

**Recommendation 3:** The Chief Research and Analytics Officer should establish, document, and implement procedures to ensure that documentation associated with internal controls and procedures around DIF's scoring methodology and model development process are reviewed and updated regularly and consistently. In addition, ensure that periodic evaluation of potential DIF alternatives are completed and adequately documented. Evaluations of alternatives should cover the composition of the scoring methodology, including all aspects and characteristics used to score returns.

> **Management's Response:** The IRS agreed with this recommendation. In its response, the IRS stated that it will develop a document that outlines the program process in a sequential timeline format to add clarity to these processes as well as adequately document efforts to evaluate potential DIF alternatives.

[21] Based on the 2018 U.S. Monitoring Report, FY 2018.

# Appendix I

## Detailed Objective, Scope, and Methodology

Our overall objective was to determine whether the IRS's systemic and nonsystemic methods and processes for scoring, selecting, and delivering tax returns of nonfarm sole proprietors with Schedules C are effective at identifying potential noncompliance. To accomplish our objective, we:

- Identified and evaluated the SB/SE Division's policies and procedures in place during Calendar Years 2013 through 2017 that control the process of Schedule C audit selection, and interviewed IRS management and subject matter experts to confirm the return selection process.
- Identified the results of research projects, CIPs, or other internal work performed that was intended to evaluate the consideration given to Schedule C tax returns, and determined whether the results were considered and incorporated, if there was a potential to improve the selection, into the IRS's inventory delivery processes.
- Obtained the FY 2019 Examination Plan and determined how the segments of the examined population are selected and balanced to represent those returns of highest risk and a representation of the overall population.
- Reviewed details and documentation and interviewed relevant management or subject matter experts associated with the systemic and nonsystemic portions of the IRS's inventory workstream associated with Schedule C tax returns. We assessed whether the methods and processes support the overall priorities and goals of the SB/SE Division.
- Identified ownership of the systemic portions of the IRS's inventory workstream (such as DIF scoring) to gain an understanding of the scoring and filtering processes. We included how the DIF assesses risk and establishes priority for Schedule C returns as they are identified for potential examination.
- Assessed whether the systemic portions of the IRS's inventory workstream supported the SB/SE Division Examination function's overall priorities and goals in the delivery of productive Schedule C examination inventory.
- Used the Individual Return Transaction File to identify the population of taxpayers filing Schedules C for TYs 2013 through 2017 that were processed during Processing Years 2014 through 2018. The tax return data were matched against the Audit Information Management System to determine whether the tax return was examined and, if such, the results of the examination.

### Performance of This Review

This review was performed with information obtained from the RAAS organization Compliance Modeling Lab located in Washington, D.C., and from SB/SE Division Examination function headquartered in Lanham, Maryland, during the period March 2019 through February 2020. We relied primarily on interviews with IRS personnel and reviews of available documentation. 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 reasonable basis for our findings and conclusions based on our audit objective.

Major contributors to the report were Matthew A. Weir, Assistant Inspector General for Audit (Compliance and Enforcement Operations); Linna Hung, Director; Glen Rhoades, Director; Curtis Kirschner, Audit Manager; Nancy Van Houten, Acting Audit Manager; Tina Fitzsimmons, Lead Auditor; Donna Saranchak, Lead Auditor; and Kim McMenamin, Senior Auditor.

## Validity and Reliability of Data From Computer-Based Systems

We performed tests to assess the reliability of data from the various tables on TIGTA's Data Center Warehouse (the Individual Return Transaction File and the Audit Information Management System). We evaluated the data by testing a judgmental sample to ensure that the data elements were accurate and provided the results of the data analysis to knowledgeable IRS officials. We determined that the data were sufficiently reliable for purposes of this report.

## 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 with Schedules C attached for examination. We evaluated these controls by interviewing management, reviewing Internal Revenue Manual guidance, reviewing reports, and analyzing closed examination data.

# Appendix II

## Studies Performed on the Discriminant Function and Potential Alternatives

| Document | Intent | Conclusion |
|---|---|---|
| IRS Research and Operations Analysis Division document, *Comparison of Various Scoring Formula Techniques* (1981) | Conducted internally and discussed the results of contractor-developed models and capability of implementing a different technique given computer processing resources. | Indicated that the IRS should expand efforts and resources into improving the current model. Also recommended that the IRS should consider outside research contracts every two or three years in hopes of finding newly developed techniques that improve upon the methods of selecting returns for examination. |
| IRS internal study, *The Selection of Returns for Audit by the IRS* (1987) | Conducted internally to provide a history of DIF development since inception. | While the study did not provide an evaluation of DIF development process, the author noted that the overall process would benefit from further research and studies of alternatives. |
| IRS internal study, *A Hybrid System: Combining DIF With CART* (1993) | Internal study to determine the feasibility of combining DIF with the Classification and Regression Trees (CART) model.[1] | While the combination of DIF and CART has the potential to increase audit yield by a small measure, this combination may lead to a higher no-change rate. Additionally, more complicated computer programming would be needed to implement a CART/DIF combination to score returns. |
| IRS Research Bulletin paper, *Neural Networks and Discriminant Function: Alternative Techniques of Selecting Tax Returns for Audit* (1994) | An internal analysis of predictive models using neural networks. A task force compared the results of DIF against neural networks. | Neural network results were inferior to DIF, and as a result, neural networks were not recommended as a replacement. Also, neural networks were found to be more expensive and difficult to implement and monitor as an IRS program, while not providing substantial additional benefits. |
| IRS internal report, *Alternative Tax Scoring Methodologies* (1999) | External alternative scoring methodologies were evaluated against a benchmark technology similar to the DIF methodology. | The alternative methodologies were able to achieve an improvement in average assessments without sacrificing the hit rate in three of four examination tax classes analyzed. Results also state that further research would be needed to determine whether these alternative methodologies work well in a practical setting where later tax year returns are scored. |
| IRS Research Conference paper, *Testing the UI-DIF Formulas* (2002) | Internal study performed at the request of the SB/SE Division to determine whether returns with unreported income could be identified, using existing DIF methodologies, outside of those already identified by computer matches. | Unreported Income scores were found to be a reasonable indication that there is a likelihood of unreported income. In particular, nonfarm business returns with high Unreported Income scores were almost always classified for audit by experts in unreported income. |

[1] CART is a statistical technique that identifies noncompliant groups of returns by variable splits.

| Document | Intent | Conclusion |
|---|---|---|
| IRS internal report, *Development of an Alternative Tax Scoring Methodology* (2003) | Follow-on to the 1999 study of how external alternative scoring methodologies compared to the current DIF. | Results suggested that a number of alternate approaches to variable selection, variable specification, and scoring can all lead to a reasonably effective means of selecting workload. Also, the IRS should consider that performance might be improved under the current DIF methodology by including some interaction terms, Winsorized variables, and/or ratio variables.[2] |
| IRS internal paper, *The DIF Methodology and Alternative Techniques for Workload Selection* (2003) | Document DIF methodology and acquaint interested IRS personnel with research by DIF's subject matter experts regarding improvements and alternatives to DIF from 1980 to the present. | CART and Probability Regression Model techniques were not recommended for implementation even though prior studies suggested the potential of these better models than discriminant analysis. However, the study resulted in recommendations for workload selection models using DIF and a contractor's formulas. |
| IRS internal document, *The Workload Selection Problem: A Literature Review* (2006) | Interpretation of DIF and a review of other studies conducted on the viability of DIF and other alternatives in resolving IRS workload selection issues. | Concluded that DIF methodology appears to be reasonably successful. One suggestion for improvement was using data from other sources provided operational or legal constraints do not preclude their use. |

*Source: IRS RAAS organization studies.*

[2] In the context of the referenced study, continuous explanatory variables were "Winsorized," meaning that the value of each variable was truncated so that all values above the 95th percentile were replaced with the value associated with the 95th percentile. For variables with negative reported values, separate truncations of the positive and negative values were performed, truncating the negative values at the 5th percentile and the positive ones at the 95th percentile. Ratio variables in the context of this study refers to ratios of line item variables.

# Appendix III

## Management's Response to the Draft Report

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

COMMISSIONER
SMALL BUSINESS/SELF-EMPLOYED DIVISION

August 25, 2020

MEMORANDUM FOR MICHAEL E. McKENNEY
DEPUTY INSPECTOR GENERAL FOR AUDIT

FROM: Eric C. Hylton Eric C. Hylton Digitally signed by Eric C. Hylton Date: 2020.08.26 09:41:46 -04'00'
Commissioner, Small Business/Self-Employed Division

SUBJECT: Draft Audit Report – Individual Returns with Large Business Losses and No Income Pose Significant Compliance Risk (Audit #201830016)

Thank you for the opportunity to review and comment on the subject draft audit report. As stated in IRS Policy Statement 4-21, the primary objective in selecting returns for examination is to promote the highest degree of taxpayer voluntary compliance. This requires us to utilize our various workload selection methodologies to identify tax returns for examination that pose the greatest risk for noncompliance and deploy our limited staffing resources to ensure we have a compliance presence within each class of taxpayers by income level.

During the audit we provided data that indicated there is adequate coverage for taxpayers filing with one Schedule C, no gross receipts, and more than $100,000 in losses. Increasing the coverage rate on this population would require the diversion of resources from other areas, thereby reducing the IRS' ability to address those respective compliance gaps. TIGTA's position that we need to improve our case selection is based on its review of the Compliance Initiative Project (CIP) returns examined in Fiscal Year (FY) 2016. This position fails to consider our subsequent implementation of revised Discriminant Function (DIF) algorithms, the Enterprise Planning Scenario Tool (EPST), and the Return Order and Delivery System (ROADS), which have improved the selection and delivery of DIF returns for examination. Failure to consider these improvements undermines the determination that a CIP is warranted. Moreover, the FY 2016 CIP did not examine a statistically valid sample of returns, so projecting its results to an entire population provides an inaccurate picture of cases with a high risk of noncompliance. There is no confidence the results could be repeated. The creation of a new strategy or workstream focused on this population is not warranted as these returns are already identified through other existing strategies that provide sufficient coverage.

2

The IRS staffing budget decreased over 15% from FY 2013 to FY 2018 and between FY 2011 to FY 2018 a hiring freeze was in effect. During this time, we faced increasing challenges, including implementation of new tax law provisions, such as the Foreign Account Tax Compliance Act (FATCA), the Affordable Care Act (ACA), and the Tax Cuts and Jobs Act (TCJA). Over the last two years we have been able to increase the hiring of enforcement personnel, but with an aging workforce and increased attrition, it will take considerable time for us to increase our enforcement staffing at any level of significance. Because of these funding and staffing limitations, we must make difficult decisions regarding priorities and the types of enforcement actions we pursue and service we provide. We must achieve balanced coverage across the various types of returns as well as geographically. Beginning in FY 2020, we have a renewed focus on the compliance of high-income taxpayers and on combating fraud.

We continuously work to improve the predictive models for workload selection. This includes improving the models developed using the DIF analytical technique as well as exploring the use of different analytical techniques and other potential enhancements. We will supplement our existing documentation of this ongoing work by creating a document that provides a timeline of the DIF development and implementation process for clarity.

Attached are our comments and proposed actions to your recommendations. If you have any questions, please contact me, or Scott Irick, Director, Examination Operations, Small Business/Self-Employed Division.

Attachment

Attachment

:

**RECOMMENDATION 1:**

The Commissioner, Small Business Self Employed Division, should conduct a national CIP or research project on those returns with at least one Schedule C attached, no gross receipts, and more than $100,000 in losses to evaluate it as a new strategy or workstream. The IRS should use the methods and results from this CIP or research project to consider other compliance gaps, past CIPs, or issues to ensure that the SB/SE Division is selecting the most productive returns for examination.

**CORRECTIVE ACTION:**

We have significant coverage of Schedule C returns within the Small Business/Self-Employed Division. Our various workload selection methodologies identify the most productive returns for examination while balancing coverage and other non-compliance objectives. We do not believe there is justification to commit significant additional resources to this population based on the non-statistical results of the 2016 CIP, which was conducted before our implementation of revised algorithms and methodologies for improving the selection of returns for examination.

**IMPLEMENTATION DATE:**

N/A

**RESPONSIBLE OFFICIAL:**

N/A

**CORRECTIVE ACTION MONITORING PLAN:**

N/A

**RECOMMENDATION 2:**

The Commissioner, Small Business Self Employed Division, should collaborate with the RAAS organization to improve any issue-based workload selection methodologies developed under Recommendation 1 by considering use of DIF in conjunction with other methods to enhance the return selection process and improve operational outcomes.

**CORRECTIVE ACTION:**

Based on our response to Recommendation 1, this recommendation is not applicable.

**IMPLEMENTATION DATE:**

N/A

**RESPONSIBLE OFFICIAL:**

N/A

2

**CORRECTIVE ACTION MONITORING PLAN:**
N/A

**RECOMMENDATION 3:**
The Chief Research and Analytics Officer should establish, document, and implement procedures to ensure that documentation associated with internal controls and procedures around DIF's scoring methodology and model development process are reviewed and updated regularly and consistently. In addition, ensure that periodic evaluation of potential DIF alternatives are completed and adequately documented. Evaluations of alternatives should cover the composition of the scoring methodology, including all aspects and characteristics used to score returns.

**CORRECTIVE ACTION:**
We have determined that the scope, content, and timing of updates to current documents is sufficient relative to the DIF development, implementation, and enhancements. However, we will develop a document that outlines the program process in a sequential timeline format to provide additional clarity to these processes. Additionally, as we evaluate potential DIF alternatives we will ensure that these efforts are adequately documented.

**IMPLEMENTATION DATE:**
March 15, 2021

**RESPONSIBLE OFFICIAL:**
Chief Research and Analytics Officer (RAAS)

**CORRECTIVE ACTION MONITORING PLAN:**
IRS will monitor this corrective action as part of our internal management system of controls.

# Appendix IV

## Abbreviations

| | |
|---|---|
| CART | Classification and Regression Trees |
| CIP | Compliance Initiative Project |
| DIF | Discriminant Function |
| FY | Fiscal Year |
| IRS | Internal Revenue Service |
| RAAS | Research, Applied Analytics, and Statistics |
| SB/SE | Small Business/Self-Employed |
| TIGTA | Treasury Inspector General for Tax Administration |
| TY | Tax Year |