



**Distribution of Tax Year 2014-2015
Individual Income Tax and Self-Employment Tax
Underreporting Tax Gap**

Distribution of Tax Year 2014-2015 Individual Income Tax and Self-Employment Tax Underreporting Tax Gap

Research, Applied Analytics & Statistics (RAAS)
Knowledge Development & Application Division (KDA)
Compliance Modeling Lab (CML)

Staff Contributors:

Drew Johns, Economist

Dan Rosenbaum, Supervisory Economist

Suggested Citation:

Internal Revenue Service

RAAS:KDA:CML

Distribution of Tax Year 2014-2015 Individual Income Tax and Self-Employment Tax
Underreporting Tax Gap

Washington, DC

December 23, 2024

Table of Contents

1	EXECUTIVE SUMMARY.....	2
2	MAIN RESULTS	3
	<i>Table 2.1: Distribution of TY 2014-2015 Individual Underreporting and Self Employment Tax Gap Sorting Taxpayers using DCE adjusted TPI and Estimates the Tax Gap using DCE Adjustments</i>	<i>5</i>
	<i>Table 2.2: Distribution of TY 2014-2015 Underreporting Tax Gap Using Various Income Measures for Tax Gap Estimation</i>	<i>6</i>
	<i>Table 2.3: Distribution of TY 2014-2015 Underreporting Tax Gap Using Various Income Measures for Sorting Taxpayers</i>	<i>6</i>
	<i>Table 2.4: Expanded Income Categories for the Distribution of TY 2014-2015 Individual Underreporting and Self Employment Tax Gap Sorting Taxpayers using DCE adjusted TPI and Estimates the Tax Gap using DCE Adjustments</i>	<i>7</i>
3	CONCLUSION	7
4	REFERENCES	10
5	DATA AND LIMITATIONS	11
5.1	NATIONAL RESEARCH PROGRAM (NRP)	11
5.2	DATA LIMITATIONS	12
5.2.1	Sampling Error	12
5.2.2	Measurement Error.....	12
5.2.3	Estimation Error.....	13
5.2.4	Coverage Error	13
6	MORE DETAIL ON DETECTION CONTROLLED ESTIMATION AND THE TAX CALCULATOR.....	15
6.1	DCE IMPLEMENTATION FOR THE TY 2014-2016 TAX GAP ESTIMATES	15
6.1.1	Grouping of Income Items for DCE Estimation.....	15
6.2	INCOME ITEMS ROUTINELY CLASSIFIED	16
6.2.1	DCE Specification for Items Routinely Classified	16
6.3	INCOME ITEMS NOT ROUTINELY CLASSIFIED	16
6.3.1	DCE Specification for Items Not Routinely Classified	17
6.4	NONCOMPLIANCE SPECIFICATION	17
6.5	DETECTION SPECIFICATION – REPORTED ROUTINELY CLASSIFIED INCOME ITEMS	18
6.6	DETECTION SPECIFICATION – UNREPORTED ROUTINELY CLASSIFIED INCOME ITEMS	18
6.7	DETECTION SPECIFICATION – NON-ROUTINELY CLASSIFIED INCOME ITEMS	19
6.8	PREDICTION FORMULA	19
6.9	SIMULATION OF UNDETECTED INCOME	20
6.9.1	Additional Tip Income Adjustments.....	20
6.10	TAX CALCULATOR	21
6.11	SELF-EMPLOYMENT TAXES	22

Distribution of Tax Year 2014-2015 Individual Income Tax and Self-Employment Tax Underreporting Tax Gap

1 Executive Summary

The IRS previously reported the tax year (TY) 2014-2016 tax gap by component, such as individual income tax vs. employment tax and underreporting vs. underpayment.¹ This report focuses on another dimension of the tax gap, **its distribution by income.**

The first step in analyzing the distribution of the tax gap by income is choosing an appropriate measure of income for sorting taxpayers. The very concept of the tax gap implies differences in the true income of taxpayers and what they report to the IRS. Should taxpayers with \$1 million in true income who report only \$20,000 count as high-income or low-income? A measure of true income, i.e., counting these taxpayers as having \$1 million, is the appropriate measure as it aligns with the measure of income on which they must pay tax. Consistent with the income measure used in the IRS's estimates of the tax gap, **the measure of true income includes reported income, income adjustments by examiners, and estimates of undetected income.**

Next, we adopt the definition of Total Positive Income (TPI),² which is the sum of all positive amounts for the various sources of income on the tax return, rather than Adjusted Gross Income (AGI) in constructing true income. AGI includes deductions, such as college expenses, retirement contributions, and losses from previous tax years, that move us away from our preferred measure of true income. **We therefore use TPI with income adjustments by examiners and estimates of undetected income as our preferred measure of true income.**

This report also presents the distribution of the tax gap using alternative income measures for sorting taxpayers. Provided that adjustments by examiners and estimates of undetected income are included, the use of TPI or AGI does not have significant effects on the results. **In contrast, measuring distributional effects using true income rather than reported income matters a great deal.**

This report **focuses on the underreporting tax gap for individual income tax and the self-employment (SE) tax.** For the TY 2014-2016 timeframe, this subset of the tax gap averaged \$331 billion per year, which is 67 percent of the average overall \$496 billion per year tax gap for the TY 2014-2016 timeframe.³

The share of the individual income tax and self-employment tax underreporting tax gap in TY 2014-2016 attributable to taxpayers with true income of \$200,000⁴ or more is much

¹ See [Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2014–2016 \(Publication 1415\)](#).

² The calculation of TPI used here is modified slightly from the IRS official TPI calculation to accommodate the data available for estimating the examiner determined and TPI adjusted for undetected income.

³ The TY 2016 NRP study was limited only to taxpayers that claimed certain credits. The final tax gap estimates used TY 2016 data only for estimates associated with those credits; all other line items were based on the TY 2014-2015 data. Consistent with that approach, this report uses only the TY 2014-2015 data. However, including TY 2016 NRP data has a negligible impact on the estimates.

⁴ \$200,000 in TY 2014-2015 is about \$260,000 in November 2023 CPIU dollars.

larger than those taxpayers' share of the population. Those high-income taxpayers account for about 31 percent of tax gap, yet only about 6% (25%) of families (income) in the Current Population Survey (CPS). Taxpayers with true income under \$50,000 account for about 21 percent of the tax gap but about 47% (15%) of families (income) in the CPS.

The estimates presented in this report are uncertain and subject to error. The tax gap estimates may be understated with respect to flow-through income (S corporation and partnership income) and income from offshore accounts because of the difficulty in detecting sophisticated forms of noncompliance by National Research Program (NRP) examinations.⁵ If examinations fail to provide adequate information about these forms of noncompliance, one might expect that there are other complex issues or types of noncompliance that are yet to be identified. To the extent that complex issues like flow-through income and offshore income are concentrated at the top of the income distribution, the challenge in measuring noncompliance on those issues could disproportionately understate the tax gap estimates at the top of the income distribution. In other words, what looks like relatively high compliance amongst those who self-report as high income might reflect low detection on complex noncompliant issues.

This report is organized as follows. Section 2 presents the main results and section 3 offers conclusions and summarizes issues discussed in the appendices. Section 4 provides references. Section 5 introduces data from the IRS's NRP examinations and discusses data limitations and challenges in estimating the tax gap. We focus on methodological issues and data limitations since comparisons across different segments of the income distribution are much more sensitive to methodological issues and data limitations than are estimates of the tax gap overall. Section 6 presents more detail on the method for estimating undetected income, tax gap estimation, and the tax calculator.

2 Main Results

We start our discussion of the main results by situating these results within the broader tax gap framework.

The tax gap is an estimate of the level of overall noncompliance with the provisions of the Internal Revenue Code (IRC) that were in effect during a particular timeframe.⁶ Tax gap estimates provide the IRS with periodic appraisals of the nature and extent of noncompliance and are used in formulating tax administration strategies. The **gross tax gap** is comprised of:

- (1) The **nonfiling gap**, which estimates the tax not paid on time by those who do not file required returns on time,
- (2) The **underreporting gap**, which estimates the difference between the tax reported on returns that were filed on time and the true tax liability for those returns, and
- (3) The **underpayment gap**, which measures the tax that was reported on returns that were filed on time but that was not paid on time.

⁵ See Guyton et al. (2023).

⁶ The word "tax" in the phrase "tax gap" is used broadly to encompass both tax and tax credits (including both nonrefundable credits and refundable credits).

The individual income tax and self-employment tax gap is the gross tax gap limited only to individual income taxes and self-employment taxes. In other words, it sets aside corporate taxes, employment taxes, and estate taxes.

The individual income tax and self-employment tax underreporting tax gap is further limited to the difference between the taxes reported on returns that were filed on time and the true tax liability for those returns. It sets aside the nonfiling gap and the underpayment gap for those taxes. The individual income tax and self-employment tax underreporting tax gap is an *estimate* of the difference between:

- (1) **An unobserved concept** – the amount of tax that *should have been reported* on timely filed tax returns, and
- (2) **An observed measure** – the amount of tax *actually reported* on timely filed tax returns.

Individual income tax and self-employment tax underreporting tax gap estimates are grounded in evidence from audit outcomes from the National Research Program (NRP). However, **even with an audit, the IRS does not necessarily observe the true amount of income given information asymmetry and the IRS carrying the burden of proof to establish unreported income.**

Income from complex financial situations, such as from large partnerships or from offshore accounts are exceptionally hard to detect. Other sources of income are substantially easier to detect, such as income sources subject to extensive third-party information reporting. In general, we expect detection of unreported income to vary with the complexity and opaqueness of the income sources.

Tax gap estimation incorporates a statistical technique called Detection Controlled Estimation (DCE) to adjust the examiner determined income for an estimated amount that is not detected by the examiner during the audit process. DCE relies on variation across examiners to infer the income that would be detected if top examiners conducted all NRP audits. DCE likely works reasonably well for most sources of income. However, it is unlikely that the DCE can fully account for undetected income whenever detection rates are suspected to be exceptionally low. To the extent that complex issues like digital assets, flow-through income, and offshore income are concentrated at the top of the income distribution and are challenging to detect for even top examiners, noncompliance could disproportionately understate the tax gap estimates at the top of the income distribution.

In other words, what looks like relatively high compliance amongst higher-income taxpayers might reflect low detection on complex noncompliance issues. **Given the uncertainty in the compliance data, especially among higher-income taxpayers, comparisons across income levels should be made cautiously.** The IRS is continuing its research to improve the methodology and completeness of the data to reduce the uncertainty in the estimates.

Table 2.1 presents our baseline results with a baseline specification which sorts taxpayers using DCE adjusted TPI and estimates the tax gap using DCE adjustments and including the self-employment tax. In that baseline specification, **the share of the underreporting tax gap is:**

- **21 percent for lower-income taxpayers** (less than \$50,000 of true income).

- **48 percent for middle-income taxpayers** (\$50,000 to \$200,000 of true income).
- **31 percent for higher-income taxpayers** (\$200,000 of true income).

Table 2.1.1 compares the distribution of the tax gap to the distribution of family income in the CPS. This is not an apples-to-apples comparison, because of differences between tax units and families, between TPI and income reported in the CPS, and because of nonfilers.

Lower-income taxpayers make up a higher percentage of the tax gap than their share of income (21% vs. 15%) but a much lower percentage than their share of families (21% vs. 47%).⁷ The tax gap of higher-income taxpayers is also higher their share of income (31% vs. 25%) but is much larger than their share of families (31% vs. 6%), although tax gap estimates for high-income taxpayers may be understated if income for sources, such as digital assets, flow-through income, and offshore income, tends to be under-detected.

Table 2.1: Distribution of TY 2014-2016 Individual Underreporting and Self-Employment Tax Gap, Sorting Taxpayers using DCE Adjusted TPI

Income Category (DCE Adjusted TPI)	Percentage of Tax Gap	Current Population Survey (CPS)	
		Percentage of Income	Percentage of Families
Less than \$50,000	21%	15%	47%
\$50,000 to \$200,000	48%	60%	47%
\$200,000 or More	31%	25%	6%

Our next few tables assess how much it matters to use different methods for estimating the tax gap or sorting taxpayers, where the top row shows our baseline distribution. We choose to assess how every one-step change from our preferred specification affects the results. We start first in Table 2.2 with the methods used for estimating the tax gap in each case while sorting taxpayers by DCE adjusted TPI. Here is what we find:

- **Using examiner determined income (row 2) in tax gap estimation rather than DCE adjusted income (row 1) shifts more of the tax gap to lower-income taxpayers,** mostly due to more of the income of lower-income taxpayers coming from wage and salary income, which results in taxes being withheld and income being reported by third parties to the IRS. Since underreporting of these and other highly visible sources of income is easier to detect, the DCE estimates tend to be lower for lower-income taxpayers. Omission of DCE adjustments in the tax gap estimation disproportionately misses noncompliance by taxpayers with higher true income.
- **Excluding the self-employment tax (row 3) in tax gap estimation shifts more of the tax gap to higher-income taxpayers.** This is partially due to higher-income taxpayers paying self-employment tax on a lower share of their self-employment income. We also

⁷ Income and family population shares are from the Current Population Survey (CPS). The CPS income definition differs from the TPI definition, but using the CPS allows for the inclusion of all households in that income range, including nonfilers. These are not apples-to-apples comparisons, but the CPS comparisons help put the tax gap estimates into perspective.

estimate that higher-income taxpayers misreport a lower percentage of their self-employment income.

Table 2.2: Distribution of TY 2014-2016 Underreporting Tax Gap, Using Alternative Tax Gap Estimates

Income Measure Used for Sorting Taxpayers	Income Measure Used for Tax Gap Estimation	Self-Employment Tax Included in Tax Gap Estimation	Underreporting Tax Gap Share		
			Less than \$50,000	\$50,000 to \$200,000	\$200,000 or More
DCE Adjusted TPI	DCE Adjusted	Yes	21%	48%	31%
DCE Adjusted TPI	Examiner Determined	Yes	29%	42%	29%
DCE Adjusted TPI	DCE Adjusted	No	20%	45%	35%

Next in Table 2.3, we assess how much using various income measures for sorting taxpayers matters. In all cases, the tax gap estimation incorporates DCE adjustments and includes the self-employment tax underreporting tax gap. Here is what we find:

- **Sorting taxpayers by examiner determined TPI (row 2) rather than DCE adjusted income (row 1) shifts more of the tax gap from higher-income to lower-income taxpayers.** Lower-income taxpayers tend to have a greater share of their income covered by third-party information reporting, such as wages and salary income. This income is easier to detect compared to low visibility income that makes up a relatively greater share of income for higher-income taxpayers. In the baseline, sorting taxpayers by DCE adjusted income moves some noncompliant taxpayers into a higher income category.
- **Sorting taxpayers by taxpayer reported TPI (row 3) rather than DCE adjusted income (row 1) shifts dramatically more of the tax gap from higher-income to lower-income taxpayers.** Taxpayers who underreport their income often fall into lower income categories relative to their true income. Sorting by taxpayer reported income presents a very misleading picture of the tax gap distribution by income.
- **Sorting taxpayers by DCE adjusted TPI (row 1) or DCE adjusted AGI (row 4) has a modest effect.** It does shift a little of the tax gap more towards lower-income taxpayers.

Table 2.3: Distribution of TY 2014-2016 Underreporting Tax Gap, Using Alternative Income Measures for Sorting Taxpayers

Income Measure Used for Sorting	Income Measure Used for Tax Gap Estimation	Self-Employment Tax Included in Tax Gap Estimation	Underreporting Tax Gap Share		
			Less than \$50,000	\$50,000 to \$200,000	\$200,000 or More
DCE Adjusted TPI	DCE Adjusted	Yes	21%	48%	31%
Examiner Determined TPI	DCE Adjusted	Yes	31%	45%	24%
Taxpayer Reported TPI	DCE Adjusted	Yes	45%	40%	15%

DCE Adjusted AGI	DCE Adjusted	Yes	22%	48%	30%
------------------	--------------	-----	-----	-----	-----

After the Table 2.2 and 2.3 comparisons of the results using different income measures for sorting and tax gap estimation, Table 2.4 circles back to the preferred income measure⁸ to provide more granularity in terms of income categories and presents a measure of the noncompliance, i.e., the net misreporting percentage (NMP), for those expanded income categories. This measure of NMP shows the ratio of net misreported income to total true income (DCE adjusted TPI).

Table 2.4 shows that this measure of NMP is lower for taxpayers with true income above \$1 million. This could be because of some combination of (a) misreporting is harder to detect for highest-income taxpayers, (b) highest-income taxpayers may have more flexibility to legally shield income from taxation, and (c) noncompliance rates are lower among the highest-income taxpayers. As argued in Guyton et al. (2023), these lower net misreporting percentages for higher-income taxpayers are circumstantial but not definitive evidence that examiner adjustments and DCE estimates of undetected income may understate the tax gap for the highest-income taxpayers due to difficulties in detecting income sources, such as digital assets, flow-through income, and offshore income. Outside of the taxpayers with true incomes above \$1 million, the net misreporting percentage does not vary a great deal, ranging between 8% and 10%.

Table 2.4: Expanded Income Categories for the Distribution of TY 2014-2016 Individual Underreporting and Self Employment Tax Gap Sorting Taxpayers using DCE adjusted TPI and Estimates the Tax Gap using DCE Adjustments

Income Category (DCE Adjusted TPI)	Tax Gap	Share of Tax Gap	Net Misreported Income (DCE Adjusted TPI)	TPI Net Misreporting Percentage (NMP)
Less than \$50,000	\$69 billion	21%	\$143 billion	8%
\$50,000 to \$100,000	\$75 billion	23%	\$250 billion	10%
\$100,000 to \$200,000	\$85 billion	26%	\$282 billion	10%
\$200,000 to \$500,000	\$59 billion	18%	\$186 billion	10%
\$500,000 to \$1 million	\$23 billion	7%	\$63 billion	9%
\$1 million or more	\$21 billion	6%	\$54 billion	4%
Total	\$331 billion	100%	\$979 billion	9%

3 Conclusion

This report presents the distribution of the individual income tax and self-employment tax underreporting tax gap by true income using a measure of true income that includes examiner adjustments and DCE estimates of undetected income. Given that unreported income makes up much of the tax gap, it is more meaningful to sort taxpayers by true income rather than reported income. Counting a taxpayer with \$1 million in true income who reports only \$20,000 as low income runs counter to the everyday understanding of what it means to be low income. Sorting by true income rather than reported income matters a great deal. Sorting by reported income

⁸ Table 2.4 uses our preferred income specification, DCE adjusted TPI for sorting and DCE adjusted income for tax gap estimation with the inclusion of the self-employment tax.

suggests larger shares of the tax gap for lower-income taxpayers and smaller shares for higher-income taxpayers.

Including examiner adjustments and DCE estimates of undetected income in the measure of true income that we use to sort taxpayers are both important. Our choice of using TPI rather than AGI for sorting taxpayers also matters, but to a much lesser degree than including examiner adjustments and DCE estimates of undetected income.

The tax gap estimate used for the distribution also matters. Factoring in DCE estimates of undetected income shifts more of the tax gap to higher-income taxpayers, whereas factoring in the self-employment tax does the reverse.

Sorting by our preferred measure of true income (DCE adjusted TPI) and including DCE estimates and the self-employment tax in tax gap estimates, we find that about half of the tax year 2014-2016 individual income tax and self-employment underreporting tax gap is due to taxpayers with true incomes between \$50,000 and \$200,000, roughly the same as their share of the population and share of overall income.

Lower-income taxpayers (those with true incomes below \$50,000) account for about a fifth of the tax gap, which is about half their share of the population but about twice their share of the overall income. Higher-income taxpayers (those with true incomes over \$200,000) account for a little less than a third of the tax gap, which is about the same as their share of overall income but three times their share of the population.

As mentioned above, these findings are sensitive to the choice of income measure used. They are also sensitive to the limitations in the data and methods used to estimate the tax gap, in many ways even more sensitive than estimates of the overall tax gap.

For example, the findings in Guyton et al. (2023) suggest that NRP audits may largely fail to detect noncompliance due to sources such as offshore and flow-through income, resulting in this noncompliance also likely not being part of DCE estimates of undetected income. This undetected income is highly concentrated among higher-income taxpayers. Consequently, even though offshore and flow-through income may not be huge components of the overall tax gap, they matter a great deal in analyses of the tax gap by income (and any related shifting of tax administration resources across income groups).

Section 5 provides more detailed discussion of the NRP data used in this analysis, as well as the limitations of these data and the methodology used to produce tax gap estimates. Section 6 goes into much more detail on DCE estimation, as well as providing more detail on tax gap estimation and the tax calculator. Like with data and methodology limitations, understanding the limitations of DCE is even more important for understanding estimates of the tax gap by income than it is for estimates of the tax gap overall. DCE is driven largely by what top examiners are able and unable to detect. If the sources of income that top examiners are unable to detect varies a great deal across income categories, then accounting for DCE is unlikely to accurately reflect differences in the tax gap across income categories.

4 References

- Bloomquist, Kim, Ed Emblom, Drew Johns, and Patrick Langetieg (2012) “Estimates of the Tax Year 2006 Individual Income Tax Underreporting Gap,” Paper Presented at the 2012 IRS and the Urban-Brookings Tax Policy Center Research Conference
- B. Erard & Associates (2005), “IRS Tax Gap Estimation: Preliminary Results of Detection Controlled Analysis,” PowerPoint presentation to Internal Revenue Service Office of Research, November 1, 2005.
- B. Erard & Associates (2006), “Preliminary Econometric Results,” Results summary report submitted to Internal Revenue Service Office of Research, January 27, 2006.
- B. Erard & Associates (2007), “Adjustment of Income Tax Underreporting Using Detection Controlled Estimation,” Final Report for IRS Contract Number TIRNO-05-D-00050 0001, November 15, 2007.
- B. Erard & Associates (2014), “Econometric Support for Developing DCE Estimates Using NRP Data for TY2006-2008,” Final Report for IRS Contract Number TIRNO-10-D-00021-D0003, March 21, 2014.
- Erard, Brian and Jonathan S. Feinstein (2012), “The Individual Income Reporting Gap: What We See and What We Don't,” IRS Research Bulletin, Publication 1500 (Rev. 2012).
- Feinstein, Jonathan S. (1990), “Detection Controlled Estimation,” *Journal of Law and Economics*, 33(1):233-276.
- Feinstein, Jonathan S. (1991), “An Econometric Analysis of Income Tax Evasion and its Detection,” *Rand Journal of Economics*, 22(1):14-35.
- Guyton, John, Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman (2023), “Tax Evasion at the Top of the Income Distribution: Theory and Evidence,” NBER Working Paper 28542, December 2023.
- Internal Revenue Service. *Estimation of the Underreporting Tax Gap for Tax Years 2008-2010: Methodology*, Research, Applied Analytics & Statistics Technical Paper, (3-2017).
- Internal Revenue Service. *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008-2010*, IRS Publication 1415 (Rev. 5-2016), Washington, D.C. 2016.
- Internal Revenue Service. *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2011-2013* IRS Publication 1415 (Rev. 9-2019), Washington, D.C. 2019.
- Internal Revenue Service. *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2014-2016* IRS Publication 1415 (Rev. 10-2022), Washington, D.C. 2022.
- Internal Revenue Service. *Estimation of the Underreporting Tax Gap for Tax Years 2014–2016: Methodology* IRS Publication 5784 (Rev. 10-2022), Washington, D.C. 2022.
- Internal Revenue Service. *Tax Gap Estimates for Tax Year 2008-2010*. April 2016.
- Internal Revenue Service. *Tax Gap for Tax Year 2006: Overview*. January 2012.
- Internal Revenue Service. *Tax Year 2006 Tax Gap Estimate: Summary of Estimation Methods*. January 2012.
- Internal Revenue Service. *Tax Gap “Map”: Tax Year 2006*. December 2011.
- Internal Revenue Service, Office of Research. *Federal Tax Compliance Research: Tax Year 2006 Tax Gap Estimation*, March 2012.

5 Data and Limitations

5.1 National Research Program (NRP)

The estimates of the individual income tax and self-employment tax underreporting tax gap are based on data from the IRS National Research Program (NRP).⁹ NRP is an examination program where returns are selected for audit (examination) in a statistical manner that allows one to draw inferences about the population from the results of those audits. The purpose of a given NRP audit is to ascertain the correctness of the return examined and determine the correct liability. From TY 2006-2015, the IRS conducted annual studies of approximately 14,000 individual income tax returns a year. Data for a given tax year generally are available for analysis purposes about three years after the returns are filed.

The TY 2016 NRP study consisted of a subset of taxpayers that claimed certain credits. The final tax gap estimates for TY 2014-2016¹⁰ only reflected the TY 2016 data in the estimates associated with those credits; all other line items were based on the TY 2014-2015 data. This report, therefore, uses only the tax gap estimates for TY 2014-2015. Using only the TY 2014-2015 NRP data has a negligible impact on the estimates.

NRP uses a process called classification to determine the type of audit for each return selected and the mandatory issues to be examined.¹¹ The classification process compares information return documents (Forms W-2, Forms 1099, etc.) with the actual tax return to identify discrepancies. It also identifies items that appear large, unusual, or questionable. Some line items on the return, typically those that cannot be verified through information returns, are always classified as mandatory to audit. In the case of simpler returns where information can be reconciled with third-party information and there appears to be a low likelihood that items are missing from the return, taxpayers are not audited and not even contacted. Returns that have only a small number of simpler issues identified in classification are routed to campus correspondence examination where the examinations can be handled through telephone calls, faxes, and traditional mail. More complicated returns are assigned to one of two types of audits that involve face-to-face interaction with an examiner: either an office audit handled by a Tax Compliance Officer (TCO) or a field audit handled by a Revenue Agent (RA) who may visit the taxpayer's place of business.

The classification process, by selecting an appropriate audit technique and set of issues, strives to reach an appropriate balance among the objectives of:

- Ensuring that taxpayers report the correct liability,
- Obtaining comprehensive and reliable information about reporting compliance, and
- Being mindful of taxpayer burden and examiner effort involved in an examination.

⁹ NRP conducts more than just individual income tax reporting compliance studies. It should be assumed for the remainder of this chapter that references to an NRP study refers to an individual reporting compliance study unless explicitly stated otherwise.

¹⁰ See [Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2014–2016 \(Publication 1415\)](#).

¹¹ Examples of issues include line items on the return, filing status, number of dependents, and whether an activity is engaged in for profit or as a hobby.

The number of mandatory issues on an NRP-selected audit typically exceeds the number of issues that would have been examined had the return been selected through the compliance risk-based processes. The audits selected through these latter programs generally are more limited in the scope of issues covered compared with those covered in the audits selected under NRP, although sometimes are worked in more depth than NRP audits. NRP audits, therefore, are more complete audits in scope, which is beneficial for ascertaining the accuracy of the return and determining the correct tax liability for most returns. However, for some complex returns, the greater depth of non-NRP audits may detect more noncompliance. Examiners of NRP audits also have the discretion to expand the audit to include non-classified issues, typically whenever information is uncovered during the course of the audit that causes the examiner to question those issues.

5.2 Data Limitations

The individual income tax underreporting tax gap estimates are subject to various types of error, like all estimates. There is sampling error, measurement error, estimation error, and coverage error.

5.2.1 Sampling Error

Since the individual income tax underreporting tax gap estimates are based on stratified random samples, there is sampling error. Sampling error occurs whenever the measure of interest for the sample differs from the true measure in the population. In particular, sampling error may arise from small sample sizes amongst the very high income (reported TPI greater than or equal to \$5 million) possibly understating the tax gap associated with those subpopulations. Irrespective of the challenges of detecting noncompliance on complex issues, if the distribution of the noncompliance is highly skewed amongst the very high income, then a small sample may not select any of the relatively small number of very large noncompliant returns.

5.2.2 Measurement Error

Measurement error with respect to the individual income tax underreporting tax gap estimates typically refers to the correctness and completeness of the examiner's determination of what should have been reported. Tax gap estimation assumes that the recommended adjustments made by the examiners are correct and appropriate, meaning that the examiner did not make adjustments that should not have been made during the examination. Tax gap estimates, however, do assume that there may be income that examiners did not detect that impacts the completeness of the examiner's determination.

In order to address the measurement error introduced by the possibility of undetected income, tax gap estimation uses the DCE methodology. The DCE methodology produces micro-level estimates that are added to the examiner recommended adjustment that then become the final data used to estimate the individual income tax underreporting tax gap. The final estimates with undetected income range from two to four times as large as estimates based solely on what the examiner determined. For data from TY 2014-2016, the estimated individual income tax underreporting tax gap is \$145 billion and \$278 billion, respectively, with and without the DCE adjustment. The appropriate methodology of adjusting the examiner determined income for the undetected amount has been a topic of ongoing investigation by the IRS. This report presents the

estimated distributions with and without the adjustments to shed light on the effect of the current DCE methodology on distribution results.

The tax gap estimates may be understated with respect to flow-through income (S corporation and partnership income) and income from offshore accounts because of the difficulty in detecting sophisticated forms of noncompliance by NRP examinations. If examinations fail to provide adequate information about these forms of noncompliance, one might expect that there are other complex issues or types of noncompliance that are yet to be identified. If top examiners detect these forms of noncompliance, it's possible that the current estimates reasonably account for misreported income from these sources. However, the complex nature of these issues and the resources required to examine them suggests that even the top examiners might not be able to consistently detect all noncompliance.

The tax gap estimates may understate the extent of the tax gap specifically at the top end of the income distribution. To the extent that complex issues like flow-through income and offshore income are concentrated at the top of the income distribution, the challenge in measuring noncompliance on those issues could disproportionately understate the tax gap estimates at the top of the income distribution. In other words, what looks like relatively high compliance amongst those who self-report as high income might reflect low detection on complex noncompliant issues.

The individual income tax underreporting tax gap estimates may not fully account for underclaims of refundable tax credits. NRP guidelines instruct examiners to be equally alert to discovering unreported income as well as allowing taxpayers any additional deduction or credits to which they are entitled, but it also allows examiners to use their judgement with respect to the scope and depth of their probes.¹² If underclaims of refundable credits are not fully accounted for, the current estimate of the individual income tax underreporting tax gap attributed to lower and middle income taxpayers would be overstated.

5.2.3 Estimation Error

The adjustments made to correct for measurement error may introduce estimation error, where the adjustment does not reflect the true measurement error. For example, the specific implementation of the DCE methodology used to adjust for measurement error could understate noncompliance at the top of the income distribution. If detection and noncompliance are heteroskedastic, for example, and the variance increases with income, then the DCE estimation could result in undetected income estimates that are biased downwards, perhaps substantially.

The implementation of DCE also assigns the conditional mean amount of undetected income using a simulation approach. The true conditional amount of undetected income would have an underlying distribution, with some amounts larger or smaller than others despite having the same values for the conditioning variables.

5.2.4 Coverage Error

¹² https://www.irs.gov/irm/part4/irm_04-022-004r#idm140393398330096

The tax gap estimates might not sufficiently cover certain issues or types of noncompliance. The question as to whether specific issues or types of noncompliance are “included” in the individual income tax underreporting tax gap estimates does not always have a simple answer. The fact that a separate tax gap estimate of a specific issue is not possible or not currently available does not necessarily mean that the tax gap related to that issue is not accounted for in the estimates. In some situations, the data are not collected at the level of detail necessary for reporting on an issue. In other situations, the issue may be rare and therefore there may not be sufficient data to provide an estimate with an acceptable level of precision. Some complex issues that are covered by the tax gap estimates may only be partially accounted for due to the earlier discussed measurement error.

Emerging issues is one area where there might be coverage error with respect to projections of the tax gap. For example, digital assets experienced significant growth, in terms of both the market capitalization and transaction volume, since the TY 2014-2016 time-period reflected in the most recent tax gap estimates. Until third-party information reporting on digital asset transactions is fully available, it is challenging to assess the amount of taxable income that is derived by a taxpayer from transactions of digital assets. Measuring the extent of noncompliance is also dependent on time to complete examinations and in developing the expertise to uncover any related noncompliance.

While it is sometimes assumed that the growth in digital assets necessarily means an equivalent growth in the tax gap, there are reasons to proceed cautiously. The holding of digital assets does not necessarily indicate the generation of income. Digital assets are treated as capital assets and subject to capital gains (or losses) when sold. At this point, it’s not clear how well the total global market capitalization that is highly volatile and based on daily sales at the margin, translates into the domestic taxable income of the associated digital assets. In addition, there likely is a substitution effect occurring to some extent whereby taxpayers who are noncompliant with respect to income generated through digital asset related activities were potentially already noncompliant with respect to other income generated activities and simply shifted their investments (and noncompliance) from those other activities to digital assets.

Illegally sourced income is generally outside the scope of the tax gap estimates, primarily because the overall goal of the government is to stop the illegal activity, not to tax it. However, some portion of the tax gap likely includes misreporting associated with illegal activities because taxpayers who engage in illegal activities potentially commingle their illegal activities with their legal activities. Although income generated by illegal activities is generally outside the scope of the tax gap estimates, it’s plausible that some illegal activity is reflected in the tax gap estimates. There also is the potential for overlap between the use of digital assets and illegally sourced income.

6 More Detail on Detection Controlled Estimation and the Tax Calculator

The IRS uses a methodology called Detection Controlled Estimation (DCE) to estimate income undetected by NRP examiners. This section provides background on the DCE methodology used for current individual income tax and self-employment tax underreporting tax gap estimates.¹³

6.1 DCE Implementation for the TY 2014–2016 Tax Gap Estimates

The specification of the DCE model depends on several factors that are driven by the nature of the data and NRP procedures. Income items that are *not* covered by significant information reporting are routinely classified when reported on the return. Income items that are covered by significant information reporting are not routinely classified. Table 6.1.1 shows which income items are routinely classified and which income items are not routinely classified.

6.1.1 Grouping of Income Items for DCE Estimation

Items Not Routinely Classified	Items Routinely Classified	
<i>Estimated Jointly</i>	<i>Estimated Jointly</i>	<i>Estimated Separately</i>
Wages and Salaries	Short-term Cap Gains	Schedule C
Interest	Long-term Cap Gains	Schedule F
Dividends	Rents and Royalties	
State and Local Tax Refunds	Part., S corp., Estate, Other	
Pensions and IRAs	Form 4797 Net Gains	
Gross Social Security	Other Income	
Unemployment		

In principle, the DCE estimation models the detection propensity of each income item given the examiner qualifications. It uses information about examiners including the type of examiner (TCO or RA), the experience of the examiner, and binary variables that take the value of 0 or 1 to indicate which examiner conducted the exam. In order to differentiate the detection capabilities of different examiners, the examiners included in the detection equation must have audited a sufficient number of returns with the income item being modeled. Typically, this requirement is 15 or more returns. The data requirements for DCE meant that some income items still needed to be grouped together for purposes of estimating the detection equation.

Table 6.1.1 shows the specific groupings of income items used for estimation. Income items that are routinely classified are modeled separately from income items that are not routinely classified. Schedule C and F income are primarily estimated independent of each other and independent of other routinely classified income items. Other routinely classified income items (capital gains, rental and royalty income, partnership and S corporation income, etc.) were

¹³ For additional background on prior DCE implementations, see Internal Revenue Service. *Estimation of the Underreporting Tax Gap for Tax Years 2014–2016: Methodology* IRS Publication 5784 (Rev. 10-2022)

estimated jointly with a common detection equation. Similarly, items that were not routinely classified are also estimated jointly with a common detection equation.¹⁴

Although detection is modeled with a common equation for a jointly estimated income group, noncompliance of each income item is modeled using separate equations and parameters. In other words, the equations and parameters for the likelihood and magnitude of noncompliance are not constrained to be identical across line items within a group while the detection equation and parameters are constrained. Additionally, because different examiners may have examined different income items, the overall average detection rates for a given line item could still vary within the group. Although separate detection equations would be preferred to the use of a common detection equation, there were not enough returns in the sample to support that estimation approach.

6.2 Income Items Routinely Classified

Items that lack significant information reporting are typically classified when reported. Detection is assumed to vary according to whether the line item was reported and whether unreported income was detected. Table 6.2.1 summarizes the four resulting categories.

6.2.1 DCE Specification for Items Routinely Classified

Income Item Reported

- 1) unreported income detected
- 2) unreported income not detected

Income Item Not Reported

- 3) unreported income detected
 - 4) unreported income not detected
-

6.3 Income Items Not Routinely Classified

For income items that are not routinely classified, if there is a discrepancy between the amount reported for the income item and the amount reported on information reporting (IRP) documents, then detection is assumed to vary with whether any unreported income was initially detected. If there is no discrepancy with IRP documents, then detection is assumed to vary first with whether the income item was classified for the examination, second on whether there was unreported income detected and third on whether income was reported on the return. Detection is further assumed to vary with whether unreported income is detected and whether the income item is reported.

¹⁴ These are the same groupings of income items that was used for the TY 2006-2008, TY 2008-2010, and TY 2011-2013 DCE estimation.

IRP Discrepancy

- 1) unreported income detected
- 2) unreported income not detected

No IRP Discrepancy and Income Item Classified

- 3) unreported income detected
- 4) unreported income not detected and income reported
- 5) unreported income not detected and income not reported

No IRP Discrepancy and Income Item Not Classified

- 6) unreported income detected
 - 7) unreported income not detected and income reported
 - 8) unreported income not detected and income not reported
-

6.4 Noncompliance Specification

The DCE methodology includes a two-part specification for modeling the noncompliance of a line item. The first noncompliance equation models the *likelihood of noncompliance* using a probit model. The second equation models the *magnitude of noncompliance* conditional on the presence of noncompliance using a log-normal regression specification.

$$P^* = \beta'_p x_p + \varepsilon_p \text{ (likelihood of noncompliance)} \quad (1)$$

$$\ln(N) = \beta'_N x_N + \varepsilon_N \text{ (magnitude of noncompliance)} \quad (2)$$

The term P^* in Equation (1) represents a latent variable describing the propensity for noncompliance. The propensity for noncompliance is assumed to depend on a set of tax return characteristics (x_p) and a random disturbance term (ε_p). The term β'_p represents a set of coefficients of the explanatory variables to be estimated. If P^* is less than zero (implying a relatively low propensity for noncompliance), then the income source is fully reported on the return and noncompliance (N) is equal to zero. If P^* is greater than zero (implying a relatively high propensity for noncompliance), then the income source is underreported on the return to some extent, meaning that N is greater than zero. In that case, the magnitude of noncompliance is determined by the second equation of the model, which relates the natural log of N to a set of explanatory variables (x_N) and an error term (ε_N). The term β'_N represents a second set of coefficients to be estimated.

This two-part specification assumes that:

1. ε_p and ε_N are independently distributed;
2. ε_p follows the standard normal distribution (mean zero and standard deviation one); and
3. ε_N is normally distributed with mean zero and standard deviation σ_N .

6.5 Detection Specification – Reported Routinely Classified Income Items

Whenever the income item is reported the DCE methodology models detection of underreported income using a probit model. Given the data requirements for DCE, some income items need to be grouped for purposes of estimating the detection equation, even when using NRP data pooled across multiple years. Because line items within a group are jointly estimated using a common detection equation, the methodology assumes that a given examiner had similar detection capabilities across all of the income items within the group. Equation (3) shows the specification of the detection equation.

$$D^* = \beta_D' x_D + \varepsilon_D \text{ (detection)} \quad (3)$$

The term D^* represents a latent variable describing the propensity of the examiner to uncover noncompliance when it is present and x_D is a set of explanatory variables. The explanatory variables include dummy variables for different examiners to exploit heterogeneity among examiners in their ability to detect noncompliance. The term β_D' represents a set of coefficients to be estimated, and ε_D is an error term assumed to follow the normal distribution with mean equal to zero and standard deviation σ_D . Let D be the detection rate representing the share of underreported income that the examiner is able to detect. The relationship between the latent variable D^* and D is then:

$$D = \begin{cases} 1 & D^* \geq 1 \\ D^* & 0 < D^* < 1 \\ 0 & D^* \leq 0 \end{cases}$$

6.6 Detection Specification – Unreported Routinely Classified Income Items

For income items that are not reported on the return, the DCE methodology assumes a two-stage detection process. In the first stage, examiners probe to determine whether unreported income is present. This stage jointly estimates the likelihood of noncompliance using Equation (1) and a detection equation that models the propensity of examiners to uncover unreported income. Equation (4) shows the specification of the first stage detection equation:

$$D_p^* = \beta_{dp}' x_{dp} + \varepsilon_{dp} \text{ (detection)} \quad (4)$$

The term D_p^* represents a latent variable describing the propensity of the examiner to uncover the presence of unreported income and x_{dp} is a set of explanatory variables. The explanatory variables include dummy variables for different examiners, similar to Equation (3). The term β_{dp}' represents a set of coefficients to be estimated, and ε_{dp} is an error term assumed to follow the standard normal distribution with mean equal to zero and standard deviation equal to one.

The second stage assumes unreported income has been detected. The second stage models noncompliance and detection with equations similar to Equation (2) and Equation (3). However, since unreported income has already been detected at this stage, detection is necessarily positive. Therefore, the relationship between D and the latent variable D^* is:

$$D = \begin{cases} 1 & D^* \geq 1 \text{ (complete detection)} \\ D^* & 0 < D^* < 1 \text{ (partial detection)} \end{cases}$$

6.7 Detection Specification – Non-Routinely Classified Income Items

Non-routinely classified income items are modeled separately for the three separate scenarios specified in Table 6.3.1. The detection specification for each scenario follows a two-stage detection specification. The first stage for each scenario models whether unreported income is present using the same specification described in Section 6.6. The first stage controls for the possibility that examiners do not always detect the presence of unreported income when it exists. The second stage models the magnitude of noncompliance and is different for each scenario. Whenever there is an IRP discrepancy, the specification is the same as in Section 6.6 and controls for the possibility of partial detection of the magnitude of unreported income. For the other two scenarios where there is not an IRP discrepancy, there are not enough observations to control for partially detected noncompliance. Therefore, Equation (2) is estimated without controlling for the possibility of undetected income.

6.8 Prediction Formula

After estimating the DCE models, the formulas for the probability of undetected noncompliance and the expected magnitude of total noncompliance (detected plus undetected) are below. The formulas are conditional on whether noncompliance was initially detected. The term A refers to the magnitude of detected noncompliance.

Probability of Undetected Noncompliance When $A > 0$:

$$\Pr(N > A | A; A > 0) = \frac{\int_0^1 \frac{1}{\sigma_N A} \phi\left(\frac{\ln(A/D) - \beta'_N x_N}{\sigma_N}\right) \frac{1}{\sigma_D} \phi\left(\frac{D - \beta'_D x_D}{\sigma_D}\right) dD}{\frac{1}{\sigma_N A} \phi\left(\frac{\ln(A) - \beta'_N x_N}{\sigma_N}\right) \Phi\left(\frac{\beta'_D x_D - 1}{\sigma_D}\right) + \int_0^1 \frac{1}{\sigma_N A} \phi\left(\frac{\ln(A/D) - \beta'_N x_N}{\sigma_N}\right) \frac{1}{\sigma_D} \phi\left(\frac{D - \beta'_D x_D}{\sigma_D}\right) dD} \quad (5)$$

Predicted Magnitude of Noncompliance [Detected (A) Plus Undetected] When Undetected Noncompliance is Present and $A > 0$:

$$E(N | A, N > A; A > 0) = \frac{\int_0^1 \frac{1}{\sigma_N D} \phi\left(\frac{\ln(A/D) - \beta'_N x_N}{\sigma_N}\right) \frac{1}{\sigma_D} \phi\left(\frac{D - \beta'_D x_D}{\sigma_D}\right) dD}{\int_0^1 \frac{1}{\sigma_N A} \phi\left(\frac{\ln(A/D) - \beta'_N x_N}{\sigma_N}\right) \frac{1}{\sigma_D} \phi\left(\frac{D - \beta'_D x_D}{\sigma_D}\right) dD} \quad (6)$$

Probability of Undetected Noncompliance When $A \leq 0$:

$$\Pr(N > 0 | A \leq 0) = \frac{\Phi(\beta'_p x_p) \left[1 - \Phi\left(\frac{\beta'_D x_D}{\sigma_D}\right) \right]}{1 - \Phi(\beta'_p x_p) \Phi\left(\frac{\beta'_D x_D}{\sigma_D}\right)} \quad (7)$$

Predicted Magnitude of Noncompliance When Undetected Noncompliance is Present and $A=0$:

$$E(N|N > 0; A \leq 0) = \exp\left(\beta'_N x_N + \frac{1}{2} \sigma_N^2\right) + A. \quad (8)$$

6.9 Simulation of Undetected Income

The DCE formula underlying the return level predictions predicts a positive probability of undetected income for most returns (though this is typically very small for returns where no unreported income was detected). Simply multiplying the predicted probability of undetected income by the predicted magnitude of undetected income would result in nearly every return receiving some positive amount of undetected income for each income item, but that would not produce a realistic distribution of undetected income. In order to have a more realistic allocation of undetected income, a simulation approach is used to apply the DCE prediction formulas. The simulation process randomly allocates undetected income for a given income item based on the probability of undetected income for that item on each return.

The specific steps of the first stage of the simulation are described below. For each return:

- Step 1:** Calculate the predicted probability of the presence of undetected income conditional on whether unreported income was detected by the examiner.
- Step 2:** Calculate the predicted magnitude of total unreported income conditional on the presence of undetected income.
- Step 3:** Draw a random number between 0 and 1.
- Step 4:** If the random number is less than or equal to the predicted probability from Step 1, allocate the predicted total (detected + undetected) amount of unreported income from Step 2. Otherwise, allocate only the detected amount of unreported income (if any).
- Step 5:** Calculate the weighted sum of predicted total unreported income from Step 4 across all returns to estimate unreported income for the population.

Steps 3 to 5 were repeated ten times for each income item to create ten sets of pooled NRP TY 2014 – 2015 NRP data with simulated undetected income.

6.9.1 Additional Tip Income Adjustments

DCE is unlikely to fully account for all undetected tip income. Since tip income is relatively concentrated in a few industries and occupations, tip income represents a relatively small amount of overall wages, salaries, and tips. However, since a significant portion of tip income is paid in cash by customers, tip income is subject to less information reporting than most wages and salaries. The lack of complete information reporting and the cash nature of tips suggest that tip income had a lower compliance rate than other wages and salaries and was harder to detect during an audit. Given the concentration of tip income and the nature of the NRP samples, DCE

estimation did not support estimates of unreported tip income. Unreported tip income was assumed to have the same noncompliance rate as the detected noncompliance rate for sole proprietor net income or loss. Thus, reported tip income was multiplied by a tax year specific adjustment factor (about 0.55) based on sole proprietor net income or loss estimates to estimate unreported tip income.

6.10 Tax Calculator

To estimate underreported taxes resulting from the underreported income at the line-item level, a tax calculator was applied to individual observations (i.e., tax returns) from the ten simulated TY 2014-2015 NRP data sets. This process provided ten underreporting tax gap estimates for each line item which were then averaged to produce the final underreporting tax gap estimate. These final line-item underreporting tax gap estimates were then summed to estimate the overall individual income tax underreporting tax gap. The specific process for estimating the underreporting tax gap for each line item using the tax calculator is described below. Essentially, the additional income for each income item was added (or subtracted) to the reported amount of income and tentative tax calculated. Then that additional income was dropped and the process repeated for the next income item.

Using the first simulated TY 2014-2015 NRP data set from the first stage of the simulation:

Income

- Step 1:** Calculate tentative tax based on reported income, deductions, exemptions, and filing status.
- Step 2:** Add net misreported wages, salaries, and tips and recalculate tentative tax.
- Step 3:** Subtract tentative tax calculated in Step 1 from tentative tax calculated in Step 2. This generates an estimate of the underreporting tax gap for wages, salaries, and tips from the first simulated data set.
- Step 4:** Remove the unreported wages, salaries, and tips added during Step 2.
- Step 5:** Repeat Steps 2 to 4 for the remaining income items.

Adjustments to Income, Deductions and Exemptions

- Step 1:** Calculate tentative tax based on reported and net misreported income, reported deductions, reported exemptions, and reported filing status.
- Step 2:** Add net misreported adjustments to income (other than the deduction for one-half of self-employment tax) and calculate tentative tax.
- Step 3:** Subtract tentative tax calculated in Step 1 from tentative tax calculated in Step 2. This generates an estimate of the underreporting tax gap for adjustments to income.

Step 4: Incrementally repeat steps 2 and 3 for the one-half of self-employment tax deduction, total deductions (itemized and standard), and the exemption amount.

Filing Status

Step 1: Calculate tentative tax based on reported and net misreported income, reported and net misreported deductions, reported and net misreported exemptions, and reported filing status.

Step 2: Calculate tentative tax based on reported and net misreported income, reported and net misreported deductions, reported and net misreported exemptions and the filing status that should have been reported.

Step 3: Subtract tentative tax calculated in Step 1 from tentative tax calculated in Step 2. This generates an estimate of the underreporting tax gap associated with errors in filing status.

Credits

Step 1: Calculate refundable and nonrefundable credits based on reported income, reported deductions, reported exemptions, and reported filing status

Step 2: Calculate refundable and nonrefundable credits based on reported and net misreported income, reported and net misreported deductions, reported and net misreported exemptions and the filing status that should have been reported.

Step 3: Subtract credits calculated in Step 2 from credits calculated in Step 1. This generates an estimate of the underreporting tax gap for total credits from the first simulated data set.

These steps are then repeated for the remaining nine simulated TY 2014-2015 NRP data sets from Stage 2 of the imputation.

6.11 Self-Employment Taxes

Self-employment taxes are required to be reported by individuals with self-employment income on individual income tax returns. The underreporting of self-employment income (primarily income reported on Schedules C and F) results in underreported self-employment taxes. Each spouse on a joint return has a separate earned income threshold above which the combined wages and self-employment income are subject to Medicare taxes but not Social Security taxes. Undetected self-employment income (Schedules C and F) was allocated to the primary taxpayer and secondary taxpayer according to each taxpayer's respective share of self-employment income as determined by the examiner. Undetected wages, salaries, and tips were allocated similarly. The tax calculator then calculated the amount of self-employment taxes that should have been reported.