Technical Paper

August 2014

Taxpayer Compliance and Sources of Error for the Earned Income Tax Credit Claimed on 2006-2008 Returns

Kara Leibel

Publication 5161 (8-2014) Catalog Number 66765W Department of the Treasury **Internal Revenue Service** www.irs.gov

This page intentionally left blank.

Taxpayer Compliance and Sources of Error for the Earned Income Tax Credit Claimed on 2006-2008 Returns

Suggested Citation:

Kara Leibel
Taxpayer Compliance and Sources of Error for the Earned Income Tax Credit Claimed on 2006-2008 Returns
Research, Analysis & Statistics Technical Paper
Internal Revenue Service
Publication 5161
Washington, DC
August 2014

This page intentionally left blank.

Table of Contents

Figures and Tables

This page intentionally left blank.

Taxpayer Compliance and Sources of Error for the Earned Income Tax Credit Claimed on 2006-2008 Returns

Executive Summary

This report presents information about the nature of errors taxpayers made when claiming the Earned Income Tax Credit (EITC) in Tax Years 2006 through 2008. This is the latest of several analyses of EITC compliance undertaken by the IRS over the years to help understand behavior and develop strategies for improving the administration of the credit. Prior to this report, the most recent analysis was conducted for Tax Year (TY) 1999 in a report commonly called the 1999 Compliance Study.

As with the earlier studies of compliance, the analysis in this report relies on audit data; in this case, the audits were conducted by IRS' National Research Program (NRP). NRP audits are like other IRS audits but provide the added benefit that they can be used for population estimates of taxpayer reporting compliance. One challenge with using audit data is that taxpayers do not always respond to or participate in the audit as required. In particular, 15 percent of EITC filers selected for an NRP audit of a TY 2006-2008 return did not participate in the audit, compared to 6 percent selected for an audit for a TY 1999 return. When this happens, the audit outcomes may not reflect their "true" eligibility for the credit.[1] To address this uncertainty, two sets of estimates are presented throughout this paper, reflecting different assumptions about the true compliance behavior of these taxpayers: the "higher" estimate assumes that audit non-participants are generally noncompliant and the "lower" estimate assumes that the true compliance of audit non-participants is the same as the compliance of otherwise similar audit participants.

We find no discernible change in the overall tendency for noncompliance between 1999 and 2006-2008. This is based on a comparison of "dollar overclaim percentages," defined as total dollars overclaimed as a percent of total dollars initially claimed for EITC (before considering IRS corrections or enforcement). In TY 2006-2008, the estimates of the overclaim percentage are 28.5 percent (lower estimate) and 39.1 percent (higher estimate). Comparable figures from the 1999 Compliance Study are 30.9 percent and 35.5 percent.

These figures and related figures in this report do not correspond directly to the EITC improper payment rate and are higher than EITC improper payment amounts. For Fiscal Years 2010 through 2013, the improper payment rate estimate averaged 24.2 percent annually.[2] Among other methodological differences, the improper payment estimates account for the effects of IRS

[1] When a taxpayer does not participate in an audit, the EITC is generally disallowed because eligibility for the credit has not been substantiated by the taxpayer as required. It is possible that audit non-participants do meet the eligibility requirements for the credit and would have otherwise been entitled to the credit but for their lack of compliance with audit procedures. "True" eligibility refers to whether or not these eligibility requirements are met, which may or may not be reflected by the audit outcome.

[2] For more information about EITC improper payments, please see the Department of the Treasury Agency Financial Reports at http://www.treasury.gov/about/budget-performance/annual-performance-plan/Pages/default.aspx.

actions to prevent or recover erroneous payments, whereas the estimates in this report do not. Because the improper payment rate accounts for the effects of enforcement, the improper payment rate is lower than the dollar overclaim percentages presented in this report. Hence, the estimates in this report do not reflect the cost of EITC errors to the Federal government.

Furthermore, these estimates do not reflect the fact that some eligible taxpayers may not have claimed the credit to which they were entitled, for instance by failing to file or to claim the credit. In addition, they do not account for offsetting errors such as may occur if one parent erroneously claimed an EITC qualifying child that should have instead been claimed by the other parent. Thus, the estimates in this report principally reflect overclaims, not net EITC errors.

While the overall tendency for noncompliance is little changed, the growth in the EITC program has led to an increase in total dollars of claims and overclaims since 1999. Averaging over returns filed for TY 2006-2008, an estimated 23.7 million taxpayers claimed an annual total of $49.3 billion in EITC, compared with 18.8 million taxpayers claiming a total of $31.3 billion in EITC in TY 1999. Total overclaims for TY 2006-2008 are estimated to be $14.0 billion (lower estimate) or $19.3 billion (higher estimate). Similar figures from the 1999 Compliance Study are $12.3 and $14.0 billion, after adjusting for inflation ($9.7 and $11.1 billion in current dollars).

The majority of taxpayers who overclaim the EITC turn out to be ineligible for the credit rather than eligible for a smaller credit amount. About 79 percent (lower) and 85 percent (higher) of amounts erroneously claimed are attributed to taxpayers who were not allowed any EITC. Still, a large fraction of the taxpayers that overclaim the EITC do so by less than $500 (44 percent according to the lower estimates and 38 percent according to the higher estimates).

The most common error made is income misreporting, occurring on two-thirds of returns with known errors; on half of returns with known errors, income misreporting is the *only* error.[3] Qualifying child errors are the second most frequent type of error, appearing on 30 percent of overclaim returns where the errors are known. Despite occurring only half as often, qualifying child errors account for by far the most dollars of overclaims. Although one cannot precisely attribute overclaim dollars to separate error types due to the occurrence of multiple errors on the same return, if qualifying child errors are considered in isolation from other errors, they account for $7.2 billion of overclaims (lower estimate) or $10.4 billion of overclaims (higher estimate). These estimates are 52 percent and 54 percent of the two respective estimates of total overclaims. If qualifying child errors are considered in conjunction with other types of errors, they may account for as little as 42 percent (lower) or 44 percent (higher) of total overclaims.

Income misreporting – and in particular self-employment income misreporting – accounts for the second highest amount of overclaim dollars, with filing status errors being the third largest contributor to overclaims. Due to provisions of the *Economic Growth and Tax Relief Reconciliation Act of 2001* (EGTRRA) that relaxed the "tiebreaker" rules, tiebreaker errors did not contribute substantially to total overclaims in 2006-2008 as they had in the last compliance study in 1999.

[3] When a taxpayer fails to participate in the audit, the credit is generally denied but the source of the error is not known.

Although qualifying child errors are responsible for the largest dollar amount of overclaims, between 73 and 87 percent of children claimed for the EITC were correctly claimed. Of the children claimed in error, the largest known error is failure to meet the residency test; roughly 75 percent of qualifying children known to have been claimed in error, or 10 percent of all children initially claimed for EITC, do not meet this requirement. The relationship test is the next most common qualifying child error: of those children known to be claimed in error, at least 20 percent fail to meet the relationship test; this is roughly 3 percent of all children claimed.

Twenty-nine percent of EITC claimants self-prepare their returns, compared with 43 percent of other taxpayers. Roughly 68 percent of EITC claimants have their returns prepared by a *paid* third party, with another 3 percent relying on free tax return preparation services offered by the IRS or IRS-sponsored programs.[4] Unenrolled return preparers are the most common type of preparer chosen by EITC claimants; 26 percent of all EITC returns, and 43 percent of paid preparer returns are prepared by an unenrolled return preparer. These are also among the most prone to error: the dollar overclaim percentage for returns prepared by unenrolled return preparers is estimated to be 33 percent (lower) or 40 percent (higher).

Although comprising only 3 percent of all returns with EITC, returns prepared by volunteers in the IRS-sponsored VITA and TCE programs have the lowest error rates. Among these returns, the dollar overclaim percentage is estimated to be 11 percent (lower estimate) or 13 percent (higher estimate).

One cannot conclude without further research whether the lower errors on volunteer-prepared returns or the higher errors on returns prepared by unenrolled return preparers stem from differences in the behavior or ability of each type of preparer, or whether they stem from selection bias – differences in the characteristics of taxpayers who seek assistance from each kind of preparer.

[4] The rate at which EITC claimants use paid preparers has declined measurably in the years since 2006-2008, so these figures should not be considered current.

Introduction

The Earned Income Tax Credit (EITC) is a refundable tax credit that supplements the earned income of low-income workers. In the late 1990s, the IRS conducted a series of studies as part of special appropriations from Congress to better understand compliance issues specific to the EITC and to aid EITC administration. These studies culminated in the IRS report, *Compliance Estimates for Earned Income Tax Credit Claimed on 1999 Returns*, known informally as the 1999 Compliance Study. In addition to providing estimates of EITC overclaims, that report was used to develop strategies for improving the administration of the credit. Since its release, it has been the authoritative source on the nature of EITC compliance.

This technical paper presents new estimates of taxpayer compliance behavior related to the EITC, using data from the IRS National Research Program's (NRP) *Individual Income Tax Reporting Compliance Study* for Tax Years (TY) 2006 through 2008. The new estimates in this paper provide the basis for an updated report titled *Compliance Estimates for the Earned Income Tax Credit Claimed on 2006-2008 Returns*. Additional discussion and detail beyond what is in the updated report is presented in this technical paper, including the statistical precision of the estimates.

The Earned Income Tax Credit

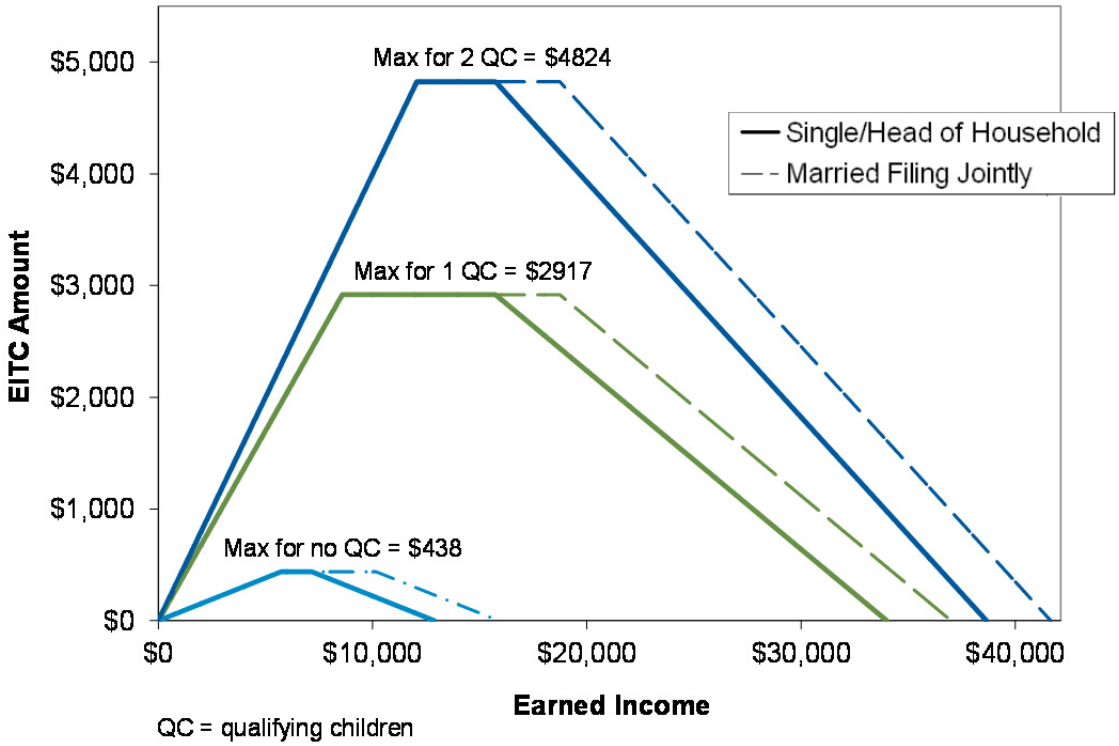
The Earned Income Tax Credit was introduced at the federal level in Tax Year (TY) 1975 as a small, refundable credit intended mainly to offset the payroll tax for low-income families. The maximum credit was $400 (roughly $1,700 in 2008 dollars) and the maximum threshold for income was $8,000 (roughly $34,500 in 2008 dollars). The credit, which was only available to taxpayers with at least one dependent, phased in at 10 percent up to $400, then began phasing out at the same rate, so that the credit decreased to $0 when earned income reached $8,000. In that first year, TY 1975, 6.2 million taxpayers claimed a total of $1.3 billion in EITC, which constituted 8 percent of all individual income tax returns for the year and totaled $5.4 billion in constant 2008 dollars.[5] Since that time, the credit has expanded steadily and has largely been re-framed as a tax incentive designed to encourage work. By TY 2008, the credit reached $51 billion and was claimed on 17 percent of all returns filed.[6]

The fundamental structure of the credit has withstood the steady expansion and is perhaps best illustrated by a trapezoidal shape, familiar by now to many who work with the EITC. Figure 1 shows the structure of the credit in TY 2008. The credit "phases in" at a fixed rate of earned

[5] IRS, *Statistics of Income--1975 Individual Income Tax Returns*, Table 3B page 69.
[6] See IRS Statistics of Income *Individual Income Tax Returns* report for TY 2008. The total number of returns claiming EITC in 2008 is 24.7 million.

income until the maximum credit amount is reached, then it remains steady over another range of income (the "plateau") before "phasing out," decreasing at a fixed rate until the credit reaches $0. While the credit was initially independent of number of children (aside from requiring at least 1 dependent), by 2008 the credit was broken out into three different "tiers" for 0, 1 and 2 or more qualifying children (a fourth tier for 3 or more children was added in TY 2009). The phase-in rates of subsidization have also changed from the original 10 percent, increasing over time to reach 34 percent for 1 child and 40 percent for 2 children. The version of the credit available to taxpayers without qualifying children, introduced in 1994, has always had a phase-in rate of 7.65 percent, which precisely offsets the employee share of Social Security and Medicare taxes. The dotted lines in Figure 1 reflect the extended plateau and phase-out ranges for married taxpayers who file jointly; these were not originally a feature of the credit but demonstrate one more aspect of the expansion of the credit over time.

Figure 1. Structure of the Earned Income Tax Credit, Tax Year 2008



EITC Legislation Since 1999

As the size and relative importance of the credit have grown over time, so have concerns about noncompliance. As just one example, a 2001 volume on the state of the federal tax system by the Joint Committee on Taxation states, "The expansion of the EITC, the large dollar amounts involved and the refundable nature of the credit have caused the EITC to be both error prone and fraud prone," (Nellen, 2001, page 212). Legislative response to this has been to simplify the rules to encourage and improve voluntary compliance. Between 1999, the year of the last compliance study, and the first year of this study, 2006, there were two major pieces of legislation concerning the EITC that might be expected to affect compliance. The *Economic*

Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA) had several EITC-related provisions that took effect in TY 2002, most of which were geared toward simplifying the credit. The *Working Families Tax Relief Act of 2004* (WFTRA) created a "uniform definition of a qualifying child" to simplify eligibility requirements for various child-related tax benefits. The relevant pieces of these two Acts are outlined below.

There were five major provisions of EGTRRA that affected the earned income credit, four of which simplified the determination of EITC and one that offered some marriage penalty relief:

- The use of "modified AGI" (modified adjusted gross income) in the determination of the credit was eliminated and the use of the standard AGI line item was restored.[7] Modified AGI increased AGI by adding back in some or all of the losses reported for capital gains, business or farm income, and from the rental of personal property (not used in a trade or business), as well as certain other losses reported on Schedule E. The use of modified AGI reduced or eliminated the credit for a number of taxpayers who were arguably not the desired beneficiaries of the EITC. Yet it added layers of complexity by requiring these additional calculations beyond those needed for the AGI concept reported on the Form 1040. Prior to EGTRRA, the calculation of modified AGI would have had to be done not only for the taxpayer, but for any other taxpayer with the same qualifying child. This is because modified AGI was integral to the "tiebreaker rules" that determined which taxpayer was entitled to claim a shared qualifying child.
- EGTRRA relaxed the tiebreaker rules so that, beginning in 2002, they only applied when more than one taxpayer actually claimed the same child. Prior to this, the tiebreaker rules dictated that only the taxpayer with the highest modified AGI was entitled to claim the child when more than one taxpayer had the same qualifying child. With the revision, taxpayers could decide amongst themselves who would claim the child. Instances of the "wrong" taxpayer claiming the child were a significant source of noncompliance in the 1999 Compliance Study. EGTRRA was expected to reduce significantly both the complexity of the tiebreaker rules and the noncompliance associated with them, although perhaps at the cost of some revenue if EITC taxpayers became strategic about who claimed the child.
- The definition of "earned income" was simplified to no longer include nontaxable forms of employee compensation.
- EITC was no longer reduced by the amount of the Alternative Minimum Tax (AMT) shown on the Form 1040.
- The final provision of EGTRRA was not about simplification but marriage penalty relief. The structure of the credit means that two unmarried individuals living together might be eligible for a certain amount of the EITC individually (although not for the same child), but if they marry and combine incomes, they could lose eligibility or become eligible for a smaller amount of the credit. To mitigate this marriage penalty, EGTRRA legislated that, beginning in TY 2002, the credit would begin phasing out at a higher income threshold for taxpayers whose filing status is married-filing-jointly. This extension was to be $1,000 higher in 2002-2004, $2,000 higher in 2005-2007, $3,000 in 2008, $5,000 higher in 2009, and indexed for inflation in subsequent years. Figure 1 shows the $3,000 extension effective in 2008. This change increases the credit for taxpayers in a certain range of income and even makes some previously ineligible now eligible. While it does not fully offset the marriage penalty, it does provide some relief, although at the cost of an added element of complexity.

[7] AGI was used in determination of the credit through tax year 1995.

At the time the *Working Families Tax Relief Act of 2004* was enacted, there was concern about the challenges taxpayers faced in negotiating the different eligibility criteria for the multiple child-related tax benefits.[8] WFTRA created a "uniform definition of a qualifying child" that went a long way toward consolidating the various eligibility criteria for these five tax benefits: the dependency exemption, Child Tax Credit (including the Additional Child Tax Credit), head-of-household filing status, EITC, and the Child and Dependent Care Credit. The new legislation defined a qualifying child as one that meets four criteria with respect to a particular taxpayer; these are informally referred to as the relationship, age, residency, and support tests.

Despite the new "uniform definition," there continue to be some differences across the various tax benefits. For example, the support test does not apply to the EITC, and the age requirements for the Child Tax Credit and the Child and Dependent Care Credit are more restrictive than the uniform definition.[9] This means that in practice, even when a child meets the uniform definition, the child can be claimed for some benefits but not others. The criteria that the *taxpayer* must meet also differ across the benefits. Thus, despite the move toward simplification under WFTRA, there is remaining complexity that may lead to some noncompliance; moreover, some new noncompliance may be introduced by the incorrect belief that the definition of qualifying child is indeed uniform or identical for all benefits. The discrepancy about the support test could mean that some taxpayers will not claim EITC for their eligible children based on the misconception that their children must meet the support test in order to be considered a qualifying child.

Another implication of the uniform definition is that taxpayers can no longer split these tax benefits by using the same child to claim different benefits on different returns (e.g., one taxpayer claims the child for EITC and another claims the child for the dependent exemption). With the uniform definition, only one taxpayer can claim a qualifying child for any of the five named tax benefits, although there are some exceptions for divorced or separated parents. This aspect of WFTRA could have had some revenue-saving effects but may have led to additional noncompliance.

There were a few other legislative changes in the period between 1999 and 2006 that were not expected to affect more than a small proportion of EITC claimants:

- For TY 2000, the definition of an "eligible foster child" was narrowed. It had previously been defined as a child whom the taxpayer cared for as his/her/their own and who lived with the taxpayer for the whole year, but beginning in TY 2000, an eligible foster child was further required to either be related to the taxpayer (sibling, step-sibling, or a descendant of these) or to have been placed with the taxpayer by an authorized placement agency.
- For TY 2001, there was a relaxation of the residency requirement for parents of kidnapped children.
- For TY 2002, the "eligible foster child" definition was again amended so that a foster child was only required to live with the taxpayer for more than half of the year rather than the whole year. This was a simplification that aligned the residency test for foster children with that for other qualifying children.

[8] See Holtzblatt and McCubbin (2003) for a discussion of the complexity of the various child-related tax benefits.
[9] See Maag (2011) for more discussion of the inconsistencies that remain across the child-related tax benefits.

- As part of WFTRA, effective beginning in TY 2004 the definition of earned income was revised to allow nontaxable combat pay to be included. The decision to do so was left up to the taxpayer since increasing income by the amount of the nontaxable pay could either increase or decrease the amount of the credit.

Another development related to new legislation since the 1999 Compliance Study is the advent of annual reporting of an EITC "improper payment rate." This measure of EITC error was developed in response to the *Improper Payments Information Act of 2002* (IPIA).[10] It is reported by IRS annually on a fiscal year basis and is included in the Department of the Treasury's Agency Financial Reports. Because the effects of IRS enforcement are now estimated annually as part of the improper payment reporting, that aspect of the analysis, which was included in the 1999 Compliance Study, has been omitted from this report. However, discussion of IRS enforcement activities is provided in the section following this one.

The dollar overclaim percentages that are presented in this report are not equivalent to the improper payment rate. The improper payment rate is estimated using a different methodology that, among other things, accounts for amounts that are recovered by IRS enforcement activities. This means the improper payment rate will be lower than the dollar overclaim percentages presented in this report, which are based on gross overclaims.

There have been some changes to the tax environment since 2008 that may have affected taxpayer compliance behavior, but any such effects will be reflected only in future studies of compliance. These changes include the following:

- Expanding the credit to a third child and increasing the income phase-out range for married-filing-jointly taxpayers, under the *American Recovery and Reinvestment Act of 2009*, effective in TY 2009;
- Establishment of new criteria governing who is allowed to claim a qualifying child when more than one person can claim the same child (creating complexity not present in the simplified tiebreaker rules under EGTRRA), under the *Fostering Connections to Success and Increasing Adoptions Act of 2008*, effective in TY 2009;
- Amending the "age test" for qualifying children, adding the requirement for the child to be younger than the taxpayer or spouse, also under the *Fostering Connections to Success and Increasing Adoptions Act of 2008*, effective in TY 2009;
- The IRS ending its practice of providing a taxpayer's "debt indicator" to financial institutions and tax return preparation firms, effective in TY 2011; this is expected to severely restrict or eliminate refund anticipation loans (but not refund anticipation checks); and
- Elimination of the Advance Earned Income Tax Credit under Public Law 111-226, effective in TY 2011.

IRS Enforcement and Activities

The IRS has made significant improvements in its administration of the EITC since 1999 and continues to devote substantial resources toward protecting and recovering revenue associated

[10] More recent legislation has revised the language and requirements of IPIA: the *Improper Payments Elimination and Recovery Act of 2010* and the *Improper Payments Elimination and Recovery Improvement Act of 2012*.

with erroneous EITC claims. For example, "pre-refund" audits have been introduced so that now more than half of EITC audits occur before the EITC portion of a refund is released to the taxpayer; the use of third-party data and updated filters have improved audit selection; and an income document matching program has been established specifically for EITC cases. In Fiscal Year (FY) 2006, roughly $2.6 billion was protected or recovered through all EITC compliance and math error programs; in FY 2008, this total was $3.7 billion.[11]

The estimates in this report will generally not reflect the impact of IRS enforcement efforts. This study focuses on characterizing *taxpayer behavior*, specifically the behavior of those claiming the EITC as it is reported on tax returns that are successfully filed and processed. IRS enforcement described above occurs after returns have been filed and accepted and therefore does not describe or directly influence taxpayer behavior.

Some IRS programs actively prevent certain submitted tax returns from being successfully filed or processed; these IRS efforts affect what aspects of taxpayer behavior are measured in this report. For instance, some errors made on electronically filed returns cause the return to be rejected. This can happen when the Social Security Number, name, or birth date of a child claimed for EITC does not match with Social Security Administration data; it can also happen when a child claimed for EITC has already been claimed by another taxpayer.[12] Other returns are not rejected up front, but are subsequently flagged as potential identity theft cases and are diverted out of the normal submission-processing pipeline. These two sets of tax returns represent potential areas of EITC compliance behavior that are not covered by the analysis in this paper, but may be worth incorporating in future studies.

To the extent that IRS enforcement and outreach activities influence compliance behavior by deterring or preventing erroneous EITC claims, those effects will be reflected in the estimates in this paper but they cannot be measured or separately identified from other factors influencing behavior.

The Data

The data used for this analysis are collected through the IRS National Research Program's (NRP) *Individual Income Tax Reporting Compliance Studies* for Tax Years (TY) 2006 through 2008, also known as the NRP 1040 Study.[13] These data are available on the IRS Compliance Data Warehouse (CDW).[14] The purpose of NRP studies is to provide information about taxpayer

[11] These figures include all tax and credit amounts protected or recovered as a result of the enforcement efforts described; they are not limited to EITC amounts. Estimates of EITC-only amounts protected or recovered on a tax year basis are constructed as part of annual improper payment reporting. For Tax Years 2006 and 2008, these are $1.6 billion and $2.1 billion, respectively.

[12] When similar errors occur on paper-filed returns, the return is not rejected. Instead, the EITC is disallowed by math error authority and the return continues through processing without the credit. The taxpayer has the opportunity to correct the error at a later time.

[13] The National Research Program conducts studies for other taxes besides the individual income tax, including the corporate income tax (Forms 1120 and 1120S) and the employment tax (Form 941). The NRP is the modern successor to the Taxpayer Compliance Measurement Program (TCMP), which conducted studies from the 1960s through 1988.

[14] The NRP data available on the CDW form the basis for the analysis, but certain additional information was obtained from the electronic case files of individual audits through the use of RGS (Report Generation Software) or

reporting compliance behavior that can be projected to the tax filing population. Using a stratified random sample design, the NRP selects a sample of returns that can be weighted to reflect the tax filing population, then conducts audits on these taxpayers. The NRP 1040 universe consists of original (not amended) income tax returns of taxpayers living in the United States who filed in the calendar year following the period of income reporting (e.g., TY 2006 returns filed in calendar year 2007). Because the study goal is to provide a comprehensive picture of tax reporting compliance, NRP audits differ from other audits conducted by the IRS (referred to here as "operational exams") in the scope of issues examined. Emphasis is placed on ensuring that the "right answer" is obtained for all line items under audit, even small dollar issues.

The NRP's first study of individual income tax returns covered TY 2001 and consisted of over 44,000 returns, including over 6,400 returns where EITC was claimed. Beginning in TY 2006, NRP shifted to a "rolling sample" design for its 1040 Studies with smaller annual samples that, when combined over three consecutive tax years, provide the same level of statistical precision for EITC-related analysis as the original one-year TY 2001 sample.

TY 2006 was also the first year that included an explicit EITC subsample. Of the 58 strata that comprise the sample design for TY 2006, 19 apply to tax returns claiming EITC.[15] This subsample was designed primarily to meet the requirements of the *Improper Payments Information Act of 2002*, which mandates the estimation of an improper payment rate with a 90 percent confidence interval of plus-or-minus 2.5 percentage points, or a 95 percent confidence interval of plus-or-minus 3 percentage points.[16] Such a sample was initially proposed in the context of a stand-alone study, but was eventually rolled into the NRP. The NRP has continued to include the EITC subsample in its annual studies, which, for the foreseeable future, has obviated the need for separate EITC compliance studies like those conducted in the 1990s.

The new NRP 1040 Studies also collect much greater detail regarding outcomes of the EITC audits compared with the TY 2001 study, which was limited to the line-item adjustments that are the bread-and-butter of NRP data. With the enhancements of the TY 2006-2008 NRP, researchers can identify the exact nature of the error (or errors) that causes the EITC to be adjusted or disallowed. This makes it possible to conduct analyses of EITC compliance at a level of detail comparable to the 1999 Compliance Study.

As noted above, the NRP 1040 Study addresses the accuracy of reported return line items on filed returns. This makes it appropriate for a study of compliance behavior of taxpayers who claimed the EITC on their originally filed return. The NRP 1040 Study does not address individuals who do not file tax returns and, as a consequence, does not collect information from which to develop estimates of underclaims that arise when EITC-eligible taxpayers do not file a

---

CEAS (Correspondence Examination Automation Support). These data were accessed for purposes of ensuring data quality and accuracy; not all resulting adjustments have been incorporated in the NRP data on CDW so there may be some differences between the officially provided NRP 1040 data and the data used for this report.

[15] A full description of the EITC strata is provided in a later section. As with the EITC 1999 Compliance Study, the NRP EITC subsample contains only taxpayers who claimed the EITC on their original return; it does not include those who first claimed the credit on an amended return. Although not a part of the current analysis, in principle the NRP data should provide the ability to analyze the set of taxpayers who first claim the credit on an amended return rather than their original return.

[16] See Appendix C to OMB Circular A-123, page 5.

tax return. Although some estimates of underclaims are presented in this report, it should be noted that these only include cases where the taxpayer claimed some positive amount of EITC on their return. They do not include cases where eligible taxpayers did not file a tax return, nor do they include cases where the taxpayer did not originally claim the credit but whose eligibility was established during the course of the audit. The IRS separately conducts research on EITC participation (not to be confused with audit participation) that addresses these kinds of excluded underclaims (Plueger and O'Hara, 2009).

Similarly, the estimates in this report do not reflect offsetting errors. For instance, when a child is found to have been claimed incorrectly in the NRP sample, in some cases there may be another taxpayer who could have correctly claimed that child but did not do so. This analysis reflects these incorrect claims but does not adjust for the fact that the credit could have been rightfully claimed by another taxpayer. The latter part of this scenario should be covered by estimates of EITC participation rather than EITC compliance and so is not addressed here.

Since both the 1999 EITC Compliance Study and the current NRP 1040 Studies were designed to provide a representative picture of compliance behavior within the EITC population, it is natural to want to compare findings from the two studies. However, there are a number of underlying differences in aspects of the studies, including differences in sample design, sample selection, and data collection methods. This means that one cannot make comparisons between the two sets of findings that are statistically definitive; that is, one cannot determine whether any apparent differences or similarities between the two studies arise from actual patterns of taxpayer behavior over time or whether they arise as a consequence of differences in the samples and methodology. Therefore, this analysis makes no attempt to quantify the statistical significance of comparisons between the two studies.

The nature of National Research Program EITC audits

When a return is selected for NRP audit, it can be examined in one of several ways. An initial review process ("classification") determines whether a return will be audited and what type of audit technique will be used. Some returns will be classified as "accepted as filed" or "accepted with adjustments," meaning that no audit actually takes place; these are generally very simple tax returns that show no signs of noncompliance based on evidence available to classifiers.[17] For EITC cases, these are typically wage earners claiming no qualifying children, low EITC, and reporting no additional income sources. The remaining returns are classified to undergo one of three types of audit: office, field or correspondence. The statutory language in IRC § 7605 governing the time and place of audits allows the IRS discretion over which type of audit is appropriate, but factors that should be taken into consideration include the complexity of the return, convenience to the taxpayer, and which type of audit lends itself to a more efficient and effective process.[18] In some cases, the actual type of audit performed differs from the classification assignment. Any figures cited in the text refer to the actual type of audit performed rather than the original classification.

[17] Note that any adjustments made during this process are not actually assessed against the taxpayer; they are recorded only for the informational purposes of NRP.
[18] See 26 CFR § 301.7605-1.

In an office audit, the taxpayer meets with an IRS tax compliance officer (TCO) at an IRS office and brings certain documentation requested ahead of time by the TCO. A field audit is one where an IRS revenue agent (RA) travels to meet with the taxpayer out in the field, typically at the taxpayer's place of business or residence – wherever the books or records to be examined are located. For a correspondence audit, the taxpayer is sent a request to submit documentation through the mail, which is reviewed by an IRS tax examiner upon receipt. This may be followed up with written requests for additional information and may occasionally involve phone contact.

In the TY 2006-2008 NRP, 94.8 percent of returns with EITC are subject to either an office audit (58.1 percent) or field audit (36.7 percent) and therefore involve a face-to-face meeting with an IRS tax auditor. A small number (1.6 percent) are accepted as filed.[19] The remaining 3.6 percent are worked by correspondence audit, with 1.0 percentage point accounted for by cases worked by standard IRS operational exam.[20]

There are arguments to be made for excluding cases worked by operational exam from any analyses of NRP data. Operational exams may not cover the same breadth of issues that would be examined in an NRP audit, nor are they subject to NRP's additional data collection requirements. This means that data available from operational exams may be less complete across a full spectrum of tax-reporting issues than data from NRP audits. Yet, because cases are selected for operational exam based on a sophisticated set of filters, as a group they represent a particularly noncompliant part of the population.[21] To illustrate this, consider that although operational exams represent only 1 percent of the TY 2006-2008 sample of EITC audits, they represent 4 percent of total overclaims in the sample. Excluding them could have measurable effects on the various population-level estimates.[22]

These tradeoffs are worth considering in the context of particular research objectives. For a study of EITC compliance, one would want to ensure that a reasonably comprehensive set of issues were audited to lead to the determination of the correct EITC amount. Although operational exams may not cover as many issues as an NRP audit, they certainly cover the EITC itself, examining qualifying child eligibility as well as income discrepancies with third-party income sources, suspicious Schedule C income, and filing status; certain other items are often covered as well, such as dependency exemptions and the Child Tax Credit. This set of issues

[19] Some of these returns have small adjustments made to them, but these are for NRP purposes only and do not affect the taxpayer's account.

[20] This is an artifact of return submission processing. NRP return selection occurs after cases are selected for pre-refund audits but without regard to the outcome of the audit selection process. This means that some cases selected into the NRP sample are already "frozen" and claimed by operational exam before the NRP can begin working them. This conflict could be solved by incorporating the NRP sample selection directly into return processing and giving it priority, as was done for the TCMP and earlier EITC Compliance Studies.

[21] Exam filters for EITC cases have become particularly successful, having been fine-tuned to the point that, in recent years, the no-change rate of audited returns with EITC ranges between 7 and 10 percent (IRS Data Books 2008-2011). In other words, over 90 percent of returns with EITC selected for audit have the credit disallowed or reduced. (In fact, the no-change rate is even lower if one considers only those cases selected for audit specifically *because* of the EITC, rather than all audits where the EITC happened to be on the return.)

[22] The bias introduced by excluding operational exams could worsen over time if the returns that are audited pre-refund become on average even more noncompliant than other returns or if a higher percent of the population undergoes pre-refund audits, or both. These things might be expected as the IRS gets increasingly good at identifying noncompliant returns and either technological advances or the growth of refundable credits make greater numbers of pre-refund audits both more feasible and cost-effective.

seems reasonably thorough for understanding EITC compliance. The information gained by including taxpayers selected for operational exam in the sample – given that they are a particularly noncompliant group – likely outweighs what is lost by foregoing the audit of a broader set of issues. Operational exam cases are therefore included in the sample used for this study.

For purposes of designing and conducting compliance studies in the future, it would be useful to explore the tradeoffs involved in using operational exams compared with NRP-type audits. For EITC cases in particular, one possibility would be to identify the NRP cases that were flagged by exam filters but not actually selected for operational exam; the outcomes of these cases, worked by NRP procedures, could be compared with the outcomes of cases that were selected and worked by operational exam. This could provide insight about whether operational exams and NRP audits yield similar results. If not, one could potentially use the findings to model how audit outcomes for the operational exams could be estimated for inclusion in analyses of NRP data.

In the TY 2006-2008 NRP, some detail remains unavailable for the operational exam cases regarding the specific nature of the errors made when EITC is reduced or disallowed. One can observe income misreporting and filing status errors, and to some degree one can infer eligibility errors, but it is not always possible to determine with certainty when errors were made with respect to the children claimed, nor is it possible to distinguish which eligibility criteria were violated or whether multiple errors occurred. The approach in this report is to presume an eligibility error was made with respect to qualifying children when no other error is observed. This affects less than one percent of the total number of overclaim cases (weighted or unweighted).

Two alternative sets of estimates

While audit data are generally considered the best data source for studies of compliance, some uncertainty arises from the fact that not all taxpayers who are selected for audit comply with the request to meet with the auditor or provide documentation about the issues under audit. These taxpayers are referred to as "audit non-participants" or simply "non-participants" throughout this paper. In prior research, terms like "no shows" or "taxpayers who did not appear for audit" have been used to describe the situation where the taxpayer does not provide any information. But because NRP audits do not have to be conducted in person, either by design or in order to accommodate the taxpayer, "no show" becomes something of a misnomer. Moreover, the term "no show" means something specific in IRS terminology and can include a taxpayer who does not provide the auditor with information during the audit but ultimately acknowledges and agrees to the auditor's proposed adjustments by signing the report (and sometimes submitting full payment). From a data collection perspective, the taxpayer's agreement with the proposed changes arguably constitutes a form of input or participation and should therefore be treated differently. The term "audit non-participant" seems broad enough to encompass any type of audit and cannot be confused with terms used internally by IRS for a slightly different purpose.

While the IRS has standard procedures for handling a lack of response from the taxpayer in terms of adjusting particular line items and overall tax liability, these adjustments do not

necessarily meet the research goals of accurately determining reporting noncompliance.[23] Some adjustments may be correct, but there is considerable uncertainty as to whether the entirety of these adjustments reflects the underlying truth for every line item. Thus, when taxpayers do not provide the auditor with any input, their audit outcome may not reflect their true circumstances. This is particularly relevant for taxpayers claiming EITC because the rate of audit non-participation is much higher for this subsample (14.6 percent, unweighted) than it is for all other taxpayers (2.9 percent, unweighted).

Various hypotheses have been offered to explain the high rate of audit non-participation among EITC claimants. Some suggest that the large amount of the credit induces taxpayers to knowingly claim the credit erroneously, and by not responding to the notice for audit, these taxpayers are implicitly acknowledging their noncompliance. Other explanations rest on (perceived or real) characteristics of EITC taxpayers: that they tend to be more transient, less educated, have poorer language skills; they may be more easily intimidated by communication from the IRS; or they may have jobs that do not offer the kind of flexibility needed to adequately respond to the audit. It should be restated that these are only hypotheses; more research is needed to determine the relative importance of these competing explanations for the high rate of audit non-participation among EITC claimants.

Given the divergent theories about what explains the audit non-participation, the analysis in this paper largely follows the approach taken by the 1999 Compliance Study, which offers two sets of estimates (for most compliance measures) based on two different assumptions about the audit non-participants. The first assumption is that audit non-participants have claimed the credit erroneously.[24] Because this assumes a high degree of noncompliance for these taxpayers, these estimates are referred to as the "higher" estimates.[25] The second assumption is that the level of compliance of audit non-participants can be approximated by otherwise similar taxpayers who do participate in the audits. Estimates based on this assumption are termed "lower" estimates, because they assume a lower level of noncompliance among the audit non-participants than under the first assumption.[26]

The "lower" and "higher" estimates are two distinct sets of estimates, rather than two ends of a range estimated concurrently or the lower and upper bounds of a statistical confidence interval. Thus, one should not interpret either set of estimates as limits or bounds. The higher estimates do not reflect the greatest possible noncompliance, nor do the lower estimates reflect the least possible noncompliance. Because each estimate is a point estimate with its own statistical confidence interval, actual values somewhat below the lower estimate or above the higher estimate may be within the confidence interval and therefore still somewhat likely. Even aside from the question of statistical precision, true noncompliance could fall below the lower

[23] Generally speaking, the procedures in the case of a taxpayer who does not respond to the notification for audit is to disallow any tax return line items that are advantageous to the taxpayer but require documentation or substantiation.
[24] In practice, whether or not the credit is considered erroneous for purposes of this analysis depends on whether the auditor fully disallowed the credit, which is generally but not always the case.
[25] The "higher" estimates are comparable to the "upper-bound" estimates in the 1999 Compliance Study.
[26] The "lower" estimates are comparable to the "lower-bound" estimates in the 1999 Compliance Study. They are constructed by attributing compliance behavior to the audit non-participants based on the behavior of the taxpayers who did participate in the audit and who share certain basic characteristics, reflected by the sample strata and the sample weights.

estimates if the "lower assumption" is incorrect and audit non-participants are in fact *more* compliant than the audit participants. At the other end of the range, because the estimates do not make any adjustments to account for income or other errors that are not detected by the auditor, true noncompliance could fall above the higher estimates if many EITC overclaims arise from these undetected errors.

The higher estimates can also be interpreted as reflecting hypothetical outcomes if the full EITC population were audited. In that case, some fraction of taxpayers would not be able or willing to participate in the audit and would have their EITC disallowed as a result. Because audit non-participation is a facet of taxpayer compliance behavior that has relevance for tax administration, the higher estimates contain information important for the IRS, even without making any assumptions about the "true" eligibility of the audit non-participants. However, gaining a better understanding of the true compliance of these taxpayers could lead to improvements in administration of the credit. One the one hand, if it can be shown that the audit non-participants are largely noncompliant, their lack of responsiveness would make them an especially cost-effective population for future enforcement. On the other hand, if audit non-participants are largely compliant but face barriers that prevent them from participating in the audit, then they are being disallowed the EITC to which they are entitled. In this circumstance, non-participants may form a population that would benefit from focused outreach efforts rather than enforcement. The differing treatments highlight an ongoing challenge of administering the EITC: balancing the goals of reducing noncompliance among those ineligible for the credit while ensuring that eligible taxpayers receive the credit.

Reliability of National Research Program EITC audit data

The estimates in this paper come from the same data source as the EITC improper payment estimates, which have been criticized publicly for relying on audit data (National Taxpayer Advocate (NTA), 2011; Wancheck and Greenstein, 2011). Misconceptions about these data have led some researchers to anticipate a "key methodological problem" in this and future IRS reports (Wancheck and Greenstein, 2011). Concerns about the reliability of NRP audit data are addressed here in order ensure that the estimates in this report are interpreted in the right context.

The criticisms leveled at the improper payment rate and the underlying NRP data stem from a fundamental mistrust about whether audits correctly determine the amount of EITC due the taxpayer. The evidence typically presented to justify this mistrust comes from a single study that evaluates results from "audit reconsiderations" (National Taxpayer Advocate, 2004).[27] The study finds that 43 percent of the sampled audit reconsiderations led to the reinstatement of some or all of the originally claimed EITC. This study also finds that a greater number of phone contacts between the taxpayer and the IRS or Taxpayer Advocate Service (TAS) representative during the audit reconsideration is associated with a higher likelihood of having EITC reinstated. The authors of the NTA study suggest that correspondence audits often may not determine the correct amount of EITC, and this may be due to inadequate communication between the examiner and the taxpayer during the audit.

[27] "Audit reconsideration" is the term used by IRS to describe the process when taxpayers contact the IRS or Taxpayer Advocate Service after their audit has closed in order to reevaluate the audit results. The NTA study analyzed a sample of 679 EITC audit reconsideration cases that closed between July 2002 and January 2003. This sample represented the 66,893 EITC audit reconsiderations for FY 2002 (NTA, 2004, page 8).

It would be a mistake to conclude from the study that 43 percent of all EITC audits are likely to have the wrong outcome. As the NTA study states, the sample it applies to is by no means representative of the full population of EITC audits due to selection bias: presumably only those taxpayers who were convinced of their true eligibility would be likely to seek audit reconsiderations, so one would expect the percent of audit reconsiderations having the credit reinstated to be quite high. There is nothing to suggest that outcomes from the larger population of audited EITC taxpayers – those who were ostensibly satisfied with the audit outcome despite potentially being denied a large dollar amount – would have anything like the same outcome as those in the NTA study.

What the NTA (2004) study findings do allow one to conclude is that, in this time period, correspondence audits as a whole incorrectly denied or reduced the credit roughly 7 percent of the time: according to the study, roughly 28,000 taxpayers had their credit reinstated through audit reconsideration in a period when audits of EITC taxpayers averaged 400,000 each year.[28] Of this 7 percent, approximately 3 percentage points can be attributed to cases where the taxpayer did not respond to any of the audit notices during the required time frame but were later determined eligible.[29] This implies that roughly 4 percent of EITC audits were cases where the taxpayer participated in the audit, yet the audit resulted in an incorrect determination for EITC.

These findings should not decrease the credibility of the NRP data used in this study. The first reason is that, as described earlier, correspondence audits account for less than 4 percent of the EITC NRP sample; so even if 4 percent of these audits result in an incorrect determination, that would describe less than 0.2 percent of the sample.[30] Instead, the vast majority (95 percent) of NRP audits involve a face-to-face meeting between the auditor and the taxpayer.[31] This meeting is often preceded and/or followed up by phone contact and, where necessary, additional meetings take place. Because the NRP audits are oriented toward generating high quality, accurate data, auditors are trained to make every accommodation to meet with taxpayers, to educate them about the necessary documentation for substantiating EITC eligibility, and to give them sufficient opportunity to obtain and supply the necessary information. Intuitively, one would expect this type of audit format to be far more likely to achieve correct determinations than one conducted

[28] In fact, the number of EITC audits in this period fluctuates considerably, but there is also variation in how long the audit reconsiderations take. The NTA study reports that the average length of time from original EITC posting to the end of the audit reconsideration is 2.7 years. Given that time frame, it seems likely that most of the audit reconsiderations were from FY 1999 (three years before the sample in FY 2002). According to the IRS *Report to Congress: IRS Tax Compliance Activities* (2003), audits of EITC returns in this time period tallied as follows: FY 1997: 365,646, FY 1998: 324,243, FY 1999: 607,308, FY 2000: 272,020, FY 2001: 413,896, FY 2002: 377,758. Although FY 1999 is an unusually high year, averaging over the period gives figures closer to 400,000.

[29] In the NTA's representative sample, 42 percent of the audit reconsiderations were cases where the taxpayer did not participate in the initial audit (a "no response" or "late response"). Since the NTA study showed that the audit non-participants had the same rate of success at EITC reinstatement as audit participants (43 percent), one can multiply the 7 percent by 0.42, giving the result that 3 percent of all EITC audits are initial audit non-participants that are later granted EITC through audit reconsiderations.

[30] Since audit non-participants are treated differently in this analysis by a process that already accounts for possible incorrect determinations, it is appropriate to use the 4 percent rate of incorrect determinations rather than the 7 percent.

[31] In practice, some of the office or field audits designed to take place in person end up being conducted by phone, mail and/or fax instead, but this is not the norm, and is only done to accommodate the taxpayer if the taxpayer cannot or will not meet in person.

through the mail, and it does not suffer from the lack of personal contact between the taxpayer and auditor that is a concern arising from the NTA study (2004).

This is not to say that NRP audits are without error. Fortunately, since NRP audits are all worked post-refund, longer timelines are involved and oftentimes the outcomes of audit reconsiderations are incorporated in the NRP data – including audit reconsiderations that may have reversed the initial determination for EITC. This will not cover the outcomes of all audit reconsiderations that arise from NRP audits, but certainly far more than pre-refund operational exams.

The NRP also conducts an extensive battery of consistency tests on the data and devotes resources to data perfection efforts. Additional evaluation of the data quality of EITC cases was conducted specifically for this report. Care was taken to ensure that the final EITC amounts could be calculated from their underlying determinants (income, qualifying children, filing status, eligibility criteria), and that the NRP-reported amounts align with IRS Master File data. Discrepancies discovered during this process were resolved by looking back at the electronic versions of case file documents to see what was recorded in the auditor's workpapers. In some cases, this led to correcting the EITC amount in the data.

For these reasons, readers should have confidence that the NRP EITC audit data used for this report are very high quality. However, no data will be perfect. One limitation is that auditors may not detect all unreported income in every audit, which may lead to an understatement of noncompliance for some taxpayers. There may be other cases in which the EITC is incorrectly denied or reduced to some taxpayers, which would lead to an overstatement of noncompliance for those taxpayers. Despite these potential data concerns, there is no evidence suggesting that the NRP data systematically either overstate or understate EITC overclaims. We believe that the lower and higher estimates together present a reasonable range for estimates of behavior for those claiming the EITC.

The analysis does not seek to address the EITC eligibility of families or individuals who do not claim the credit. One implication of this is that, when a child is erroneously claimed by one taxpayer but could have been claimed legitimately by another, the erroneous claim will be captured in this report, but the fact that the child could have been correctly claimed by another taxpayer is not.

Summary of Tax Years 2006-2008 NRP EITC sample

The combined size of the EITC sample from the TY 2006-2008 study used in this analysis is 7,635 returns. This includes audit non-participants and operational exam cases, but excludes 241 other cases that were originally selected into the sample but for which no audit took place.[32] Table A provides some information about the makeup of the sample across years by audit type and audit participation.

[32] A case selected into the sample may be formally excluded if it meets certain criteria determined by the NRP. Of the 241 cases excluded from the sample, just one is excluded because the audit had not yet closed as of the latest release of NRP data.

Table A. Summary of NRP EITC Sample in Tax Years 2006-2008
Sample size, audit types, and audit non-participation by tax year

	2006	2007	2008	Total
Sample Size				
Original sample size	2,221	2,698	2,957	7,876
Number of excluded returns	46	62	133	241
Total in sample	2,175	2,636	2,824	7,635
Audit Type				
No audit - accepted as filed or with adjustments	17	61	46	124
Operational exam	35	18	26	79
Field audit (RA)	727	992	1,080	2,799
Office audit (TCO)	1,396	1,565	1,475	4,436
Correspondence audit	0	**	197	**197
Audit Participation				
Number of audit non-participants	294	419	403	1,116
Rate of audit non-participation	13.5%	15.9%	14.3%	14.6%

**Data combined with previous row to avoid disclosure of information for specific taxpayers.

Although correspondence audits remain a low percentage of NRP audits overall, the 197 cases in TY 2008 are indicative of a trend in which NRP audits have slowly been shifting into more correspondence audits over time. To date there is no evidence that this affects the quality of the outcomes one way or the other, but for reasons already discussed it seems quite plausible that this may have an effect.[33] Certainly the IRS has found that, for returns claiming EITC, correspondence examination is expedient and efficient for protecting revenue, but from a research perspective, the cost of moving to correspondence audits may be higher in terms of information lost. The NRP itself provides a potential data set for testing this hypothesis, although one must be cognizant of the non-random classification process that NRP uses to assign cases to audit type and address this problem econometrically.

Sample weighting and related methodological considerations

Sample weights were constructed for this analysis to ensure the estimates are representative of the EITC population. These weights differ from weights provided by NRP, although they were created by a process similar to the one outlined in *Weighting Documentation: IRS-1040 TY 2008*.[34] The steps involved are summarized here. The step for non-response adjustment was excluded, for reasons discussed below.

[33] The National Taxpayer Advocate suggests that differences in non-response rates offer some evidence that the type of audit "drastically affects the outcome," citing a non-response rate from NRP audits as 15 percent (also reported here) and a rate of 70 percent from operational correspondence audits (National Taxpayer Advocate, 2011, pp 298, 310). Yet this disparity cannot inform the quality of audit outcomes because these non-response rates apply to very different underlying populations. The 70 percent figure comes from the population of EITC returns selected for audit, which at the time of selection exhibit certain characteristics associated with a high probability of noncompliance. In contrast, the 15 percent figure comes from a representative random sample of all returns claiming EITC, which ought to be a far more compliant population. Since the effects of the type of audit cannot be disentangled from the effects of the behavior of the population under study, this comparison of non-response rates does not provide convincing evidence that the nature of the audit affects the quality of the outcome.

[34] The NRP Data Dictionary on the Compliance Data Warehouse contains the weighting documentation.

Treatment of audit non-participants and operational exam cases

One of the considerations in creating sample weights is the treatment of audit non-participants. Because the audit results for those taxpayers are not based on an examination of books and records or an interview with the taxpayer, the audit results may not reflect the taxpayer's actual circumstances. In this sense, there are "missing" values for the audit-determined amounts (although these audits are not incomplete from the perspective of the IRS). One possible way to address this would be to treat the audit non-participants the way "non-response" cases are often treated in the household or individual survey literature: assign those cases weights of zero, effectively removing them from the sample, and perform a non-response adjustment to the weights of the remaining cases. This is the approach used in creating the NRP final weights, for example. However, when the non-response behavior itself is of interest, this approach is not appropriate. Furthermore, this method effectively entrenches the idea that behavior of audit non-participants can be approximated by that of participants, making the "lower" assumption the de facto assumption for all subsequent weighted analyses. This may be reasonable for some issues or taxpayers where the rate of audit non-participation is low or where there is strong justification for the "lower" assumption. This is not the case for EITC audit non-participants, who are a proportionately large subpopulation about whom little is known. It seems preferable to acknowledge that there is considerable uncertainty introduced by these audit non-participants, rather than to wave the uncertainty away. Maintaining the audit non-participants in the analysis and assigning them positive sample weights was the approach taken by the 1999 Compliance Study, and it is the approach taken here.

One might also consider excluding operational exam cases from the sample, for reasons already discussed. However, as previously argued, for purposes of studying EITC compliance, it is preferable to keep these cases in the sample. Therefore these cases are also given a positive sample weight and included in the analysis.

Creation of the sample weights

Weights were created separately for each of the three tax years, following the process described here.

Base weights

First, base weights were calculated to reflect the probability of selection. For all but a handful of cases, base weights are equal to the population size of the sample stratum divided by its sample size. The exceptions are cases where the taxpayer filed as single or head-of-household, but whose correct filing status was married-filing-jointly, as evidenced by the choice to file jointly with a spouse as a result of the audit. The weights for these cases are adjusted to account for the increased chance of selection into the NRP sample.

Adjusting NRP EITC Strata Definitions

Because the number of sample cases in certain strata was quite small, some strata were collapsed. The following table shows the definitions for the 19 strata that apply to taxpayers claiming EITC in the TY 2006 NRP sample design. The EITC amounts used to define the sample strata are the maximum credit amounts for taxpayers with no children, one child and two

children in TY 2006. Those amounts are adjusted each year to be consistent with the current EITC parameters.[35] The adjustments to the sample strata that were made when creating the sample weights are noted in the final column of Table B. It was decided to collapse strata across similar amounts of EITC claimed, because, in comparison with filing status and presence of Schedule C income, there is less reason to believe that the amount of EITC claimed necessarily distinguishes types of taxpayers. For example, a low amount of EITC claimed could reflect claiming any number of children as well as income at either the lowest or highest ends of the eligible range.

Raking

The next step is to rake the sample to population totals along several dimensions. Eight variables were used for raking in addition to the sample strata. The variables used for raking were loosely the same as the eight described in NRP's weighting documentation for TY 2006: indicator for a refund anticipation loan or check (RAL or RAC), return preparer indicator, collapsed return preparer indicator, age category, filing status/gender, form type (1040, 1040A or 1040EZ), number of EITC qualifying children claimed, and census division.[36] Population totals for these were tallied from the appropriate tables on the Compliance Data Warehouse.[37]

Table B. NRP EITC Strata Definitions and Adjustments In the Sample Weighting Process, NRP TY 2006

Sample Code	Schedule C income	EITC amount claimed	Filing status	Adjustment
27010	None	EITC <= $412	Single	
27011	None	$412 < EITC <= $2747	Single	
27012	None	$2747 < EITC <= $4536	Single	
27020	None	EITC <= $412	Married filing jointly	
27021	None	$412 < EITC <= $2747	Married filing jointly	
27022	None	$2747 < EITC <= $4536	Married filing jointly	
27030	None	EITC <= $412	Head of household	
27031	None	$412 < EITC <= $2747	Head of household	
27032	None	$2747 < EITC <= $4536	Head of household	
27101	Negative	Any	Any	
27110	Positive	EITC <= $412	Single	
27111	Positive	$412 < EITC <= $2747	Single	Collapsed with 27112
27112	Positive	$2747 < EITC <= $4536	Single	Collapsed with 27111
27120	Positive	EITC <= $412	Married filing jointly	
27121	Positive	$412 < EITC <= $2747	Married filing jointly	
27122	Positive	$2747 < EITC <= $4536	Married filing jointly	
27130	Positive	EITC <= $412	Head of household	Collapsed with 27131
27131	Positive	$412 < EITC <= $2747	Head of household	Collapsed with 27130
27132	Positive	$2747 < EITC <= $4536	Head of household	

[35] In TY 2007, these parameters are $428, $2853, and $4716; in TY 2008, these are $438, $2917, and $4824. In actuality, the EITC dollar amounts used for TY 2008 sample selection were rounded figures rather than the actual amounts, due to an error in programming. Because many taxpayers report income that puts them in the plateau region of the EITC (and thus eligible for precisely the maximum credit for 0, 1, or 2 children), the rounding of those thresholds had some effect on the relative representation of each stratum. Specifically, 183 cases, or 7 percent of the 2008 sample, were assigned to the "wrong" strata. To correct for this, sample codes were revised to reflect the *intended* strata given the claimed EITC amount; the creation of the sample weights also reflected the intended strata and used appropriate population totals.

[36] Only documentation for the TY 2008 weights is (currently) available on the IRS Compliance Data Warehouse.

[37] The tables used to create population totals for these characteristics are the IRTF_ENTITY and IRTF_F1040.

Replication

For purposes of variance estimation, 100 sets of replicate weights were created from the base weights using the JK2 method. JK2 is a version of the jackknife replication method that can be applied when there are two primary sampling units (PSUs) per strata; each set of replicate weights (or "replicate") is created by successively dropping one of the PSUs from each strata and reweighting the remaining observations. With the JK2, one ends up with the same number of replicates as there are strata. In this case, it was necessary to first create the appropriate format for the data by dividing the sample into the desired number of "pseudo strata" – in this case, 100 – and randomly assigning each observation in each stratum to one of two primary sampling units. The creation of the replicates then follows the JK2 method described. Each set of replicate weights was also raked to population totals.

Statistical precision of the estimates

Standard errors are calculated for all estimates using these replicate weights in order to provide information about the statistical precision of the estimates. In most cases, the sample size is large enough so that consideration of the standard errors does not affect the conclusions that can be drawn.

Combining across years

The method for combining the data into one representative year is simple: the sample weight for each return is divided by three – the number of annual samples being combined. This is an accepted approach, taken, for instance, by the American Community Survey for its multi-year estimates to produce averages over the period of study (Census, 2009, Page 11-16). This methodology may not be appropriate for certain kinds of analyses, and in particular may need to be revisited or modified for analyses in which the relevant tax law changes significantly between tax years in the sample. Fortunately, the tax law relevant to the EITC remained largely the same across these three years, with the exception of adjustments for inflation. The only other exception is that the extended phase-out range for married-filing-jointly taxpayers increased from being $2,000 beyond those for single and head-of-household filers in 2006 and 2007 to being $3,000 higher in 2008. Since this is a fairly subtle change, it does not preclude estimating averages across the three years by combining the annual samples and reducing the sample weights to one-third their original value. The raked weights were also divided by three so that together they produce population-level estimates for the three-year combined sample.

Adjusting for inflation

In order to have measures of dollar amounts and overclaims that are consistent across the three years, dollar figures have been adjusted for inflation to reflect constant 2008 dollars, the most recent year of the study. The index used to adjust for inflation is the same one used to make cost-of-living adjustments to the parameters defining the EITC. This is legislated to be the average Consumer Price Index for All Urban Consumers (CPI-U) from the 12-month period ending on August 31 of the calendar year preceding the tax year in question.[38] Accordingly, the

[38] See 26 USC § 32 (j) (2012).

index for TY 2006 is the average CPI-U from September 2004 through August 2005 and the indexes for the other years are calculated in a similar manner.

The next section presents the results of the analysis.

Results

The first two tables in this section provide summary measures of compliance behavior – the first is for the full EITC population and the second breaks information down by the number of children initially claimed. The subsequent analysis shows how returns and overclaim dollars are distributed by type of error, with a particular focus on income misreporting and qualifying child errors. Next, estimates of how many overclaim dollars are associated with each error type are presented, followed by a closer look at qualifying child errors. The section concludes with two tables showing what types of return preparer are chosen by EITC claimants and how the frequency and magnitude of EITC errors vary by type of return preparer.

Overview of compliance

Table 1 provides an overview of population estimates of EITC compliance, averaged across TY 2006-2008. Lower and higher estimates are presented to reflect the different assumptions about the audit non-participants. Total overclaims are estimated to be $14.0 billion (lower estimate) and $19.3 billion (higher estimate) in 2008 dollars.[39] Comparable figures from the 1999 Compliance Study, after adjusting for inflation, are $12.3 and $14.0 billion.[40]

Some of the increase in these estimates can be explained by the growth in the program during this time period: between 1999 and (the average of) 2006-2008, the number of taxpayers claiming EITC increased from roughly 19 million to almost 24 million. This expansion has been driven by a combination of population growth, expanded eligibility through the extended phase-out ranges for married-filing-jointly taxpayers, and greater take-up of the EITC by eligible taxpayers. The latter of these has been documented by collaborative research between IRS and Census (Plueger and O'Hara, 2009).[41] These changes would all be expected to push up the total overclaim amount even if the rate of noncompliance were unchanged.

The dollar overclaim percentage – defined as total overclaims divided by total EITC claims – can provide a good indication of whether compliance has changed in relative terms. In the 1999 Compliance Study, the "lower-bound" and "upper-bound" estimates, which are conceptually similar to the lower and higher estimates in this paper, were 30.9 percent and 35.5 percent,

[39] Recall that these are not the same as annual EITC improper payment amounts.
[40] These were reported to be $9.7 and $11.1 billion in current dollars (Compliance Study, 1999, Table 1).
[41] In fact, Plueger and O'Hara (2009) pre-dates the more recent analyses that show increased take-up of the EITC, but it provides the most recent publicly available discussion of the research methods and offers a foundation for the analyses that followed. The paper by Plueger and O'Hara (2009) compares an analysis of TY 2005 data with analyses conducted previously by researchers using other methods for TY 1990 and TY 1999, and concludes that participation in EITC remained fairly stable over that period. Subsequent (unpublished) analyses for more recent tax years have shown the increase in take-up. According to IRS, "Our research studies on participation, completed with the U.S. Census Bureau, show that EITC participation increased from 75–77 percent for tax year 2005 to 78–80 percent for tax year 2008," (National Taxpayer Advocate, 2011, page 307).

respectively, with a gap of 4.6 percentage points between the estimates. The current figures for the lower and higher estimates are 28.5 percent and 39.1 percent, respectively, which are 10.6 percentage points apart. With movement at both ends of the gap (i.e., the lower overclaim percentage is lower and the higher overclaim percentage is higher) we cannot conclude that overall compliance has changed between the two studies.

Table 1. Summary of EITC Compliance Estimates
Weighted Population Estimates, Annual Average, NRP TY 2006-2008
(Dollar amounts in billions of constant 2008 dollars)

Higher estimates				
	Overclaim returns	**Correct returns**	**Underclaim returns[1]**	**Total**
Number of returns (millions)	11.9	10.4	1.4	23.7
	0.1	0.1	0.1	--
Percent of total returns	50%	44%	6%	100.0%
	1%	1%	0%	--
Amount claimed	$26.0	$20.9	$2.3	$49.3
	0.4	0.4	0.2	0.2
Correct amount	$6.7	$20.9	$2.8	$30.5
	0.3	0.4	0.2	0.4
Amount overclaimed	$19.3	$0.0	$0.0	$19.3
	0.4	--	--	0.4
Amount underclaimed	$0.0	$0.0	$0.5	$0.5
	--	--	0.1	0.1
Dollar overclaim percentage[2]	--	--	--	39.1%
	--	--	--	0.8%
Lower estimates				
	Overclaim returns	**Correct returns**	**Underclaim returns[1]**	**Total**
Number of returns (millions)	10.1	12.1	1.6	23.7
	0.2	0.2	0.1	--
Percent of total returns	43%	51%	7%	100.0%
	1%	1%	0%	--
Amount claimed	$21.8	$24.7	$2.7	$49.3
	0.4	0.5	0.2	0.2
Correct amount	$7.8	$24.7	$3.4	$35.8
	0.3	0.5	0.2	0.4
Amount overclaimed	$14.0	$0.0	$0.0	$14.0
	0.4	--	--	0.4
Amount underclaimed	$0.0	$0.0	$0.6	$0.6
	--	--	0.1	0.1
Dollar overclaim percentage[2]	--	--	--	28.5%
	--	--	--	0.7%

Note: Figures may not sum due to rounding. Standard errors are presented below estimates.

1 Underclaim returns are limited to returns where EITC was initially claimed by the taxpayer on his/her filed return, consistent with the definition of underclaim returns used in the 1999 Compliance Study. This excludes returns where EITC was not claimed, even if the taxpayer was found to be eligible for the credit during the audit.

2 The dollar overclaim percentage is not the same as the improper payment rate, which is calculated on an annual, fiscal year basis by a different methodology and accounts for amounts that are recovered by IRS enforcement activities.

Mechanically, the widening of the gap between the lower and higher estimates is caused by the higher rate of audit non-participation in the more recent NRP EITC sample: 15.5 percent of the weighted EITC population compared with 6.6 percent in the 1999 study.[42] With greater audit non-participation comes greater uncertainty, which leads to a larger gap between the lower and higher estimates.

Clearly the assumptions one makes about the compliance behavior of audit non-participants are crucial for any interpretation of these percentages, and the higher the rate of non-participation, the more uncertain the results. Learning more about this population of taxpayers could potentially reduce the uncertainty surrounding the dollar overclaim percentage and other estimates of EITC compliance. It could also provide additional benefits for EITC administration. If it can be shown that these are largely noncompliant taxpayers, their lack of responsiveness would make them an especially cost-effective population for future enforcement. If instead these taxpayers face barriers that prevent them from participating in the audit, they form a population that would benefit from targeted outreach efforts rather than enforcement.

Tables 2a and 2b break out the overclaims and underclaims of taxpayers by the number of qualifying children that were initially claimed. Overclaims are also separated into whether the taxpayer was found to be ineligible for the credit or eligible for a smaller credit. Both higher and lower estimates are presented.

As shown in Table 2a, the most overclaim dollars are associated with taxpayers claiming two children ($8.4 billion lower estimate, $11.4 billion higher estimate), but this appears to be due to the larger credit available with two children rather than a greater tendency toward noncompliance: the dollar overclaim percentage does not appear to vary by number of children claimed.

Table 2b indicates that most overclaim dollars are attributed to taxpayers who are in fact ineligible for the credit. According to the higher estimate, ineligible taxpayers account for 85 percent of total overclaim dollars, with just 15 percent due to those who were eligible for a smaller credit amount. The comparable figures for the lower estimate are 79 percent and 21 percent.

[42] The unweighted audit non-participation rates are 14.6 percent and 5.5 percent, respectively.

Table 2a. EITC Compliance Estimates by Number of Qualifying Children Claimed: Dollar Amounts Reported vs. Amounts That Should Have Been Reported, Weighted Population Estimates, Annual Average, NRP TY 2006-2008 (Billions of constant 2008 dollars)

	Number of Qualifying Children Claimed						Total	
	None		One		Two			
	Reported	Should Have Reported	Reported	Should Have Reported	Reported	Should Have Reported	Reported	Should Have Reported
	Higher estimates							
EITC Correct	$0.6	$0.6	$8.2	$8.2	$12.2	$12.2	$20.9	$20.9
	0.0		0.2		0.4		0.4	
EITC Overclaim	$0.6	$0.1	$9.1	$1.8	$16.2	$4.8	$26.0	$6.7
	0.0	0.0	0.2	0.1	0.4	0.2	0.4	0.3
Taxpayers ineligible for credit	*$0.4*	*$0.0*	*$6.6*	*$0.0*	*$9.5*	*$0.0*	*$16.4*	*$0.0*
	0.0	*0.0*	*0.4*	*0.0*	*0.6*	*0.0*	*0.8*	*0.0*
Taxpayers eligible for smaller credit	*$0.2*	*$0.1*	*$2.6*	*$1.8*	*$6.7*	*$4.8*	*$9.5*	*$6.7*
	0.0	*0.0*	*0.1*	*0.1*	*0.3*	*0.2*	*0.3*	*0.3*
EITC Underclaim	$0.1	$0.2	$0.8	$1.0	$1.4	$1.7	$2.3	$2.8
	0.0	0.0	0.1	0.1	0.1	0.2	0.2	0.2
Total EITC	$1.3	$0.9	$18.1	$11.0	$29.8	$18.7	$49.3	$30.5
	0.0	0.0	0.2	0.2	0.1	0.3	0.2	0.4
Total amount overclaimed		$0.5		$7.4		$11.4		$19.3
		0.0		0.2		0.3		0.4
	Lower estimates							
EITC Correct	$0.7	$0.7	$9.8	$9.8	$14.2	$14.2	$24.7	$24.7
	0.0	0.0	0.3	0.3	0.4	0.4	0.5	0.5
EITC Overclaim	$0.5	$0.1	$7.4	$2.1	$13.9	$5.5	$21.8	$7.8
	0.0	0.0	0.2	0.1	0.4	0.3	0.4	0.3
Taxpayers ineligible for credit	*$0.3*	*$0.0*	*$4.5*	*$0.0*	*$6.2*	*$0.0*	*$11.0*	*$0.0*
	0.0	*0.0*	*0.2*	*0.0*	*0.3*	*0.0*	*0.4*	*0.0*
Taxpayers eligible for smaller credit	*$0.2*	*$0.1*	*$2.9*	*$2.1*	*$7.6*	*$5.5*	*$10.8*	*$7.8*
	0.0	*0.0*	*0.2*	*0.1*	*0.3*	*0.3*	*0.4*	*0.3*
EITC Underclaim	$0.1	$0.2	$0.9	$1.2	$1.7	$2.0	$2.7	$3.4
	0.0	0.0	0.1	0.1	0.2	0.2	0.2	0.2
Total EITC	$1.3	$1.0	$18.1	$13.1	$29.8	$21.7	$49.3	$35.8
	0.0	0.0	0.2	0.2	0.1	0.3	0.2	0.4
Total amount overclaimed		$0.4		$5.2		$8.4		$14.0
		0.0		0.2		0.3		0.4

Notes: Figures may not sum due to rounding. Standard errors are presented below estimates.

Table 2b. EITC Compliance Estimates by Number of Qualifying Children Claimed: Dollar Overclaim Percentages and Distribution by Taxpayer Eligibility Weighted Population Estimates, Annual Average, NRP TY 2006-2008
(Constant 2008 dollars)

	Number of Qualifying Children Claimed			Total
	None	One	Two	
	Higher estimates			
Total amount overclaimed (billions)	$0.5	$7.4	$11.4	$19.3
Percent of overclaim dollars attributable to:				
Taxpayers ineligible for credit	85%	89%	83%	85%
	2%	2%	2%	2%
Taxpayers eligible for smaller credit	15%	11%	17%	15%
	1%	1%	1%	1%
Total	100%	100%	100%	100%
Dollar overclaim percentage[1]	37.8%	40.6%	38.3%	39.1%
	1.3%	1.2%	1.1%	0.8%
	Lower estimates			
Total amount overclaimed (billions)	$0.4	$5.2	$8.4	$14.0
Percent of overclaim dollars attributable to:				
Taxpayers ineligible for credit	83%	85%	74%	79%
	2%	1%	2%	1%
Taxpayers eligible for smaller credit	17%	15%	26%	21%
	2%	1%	2%	1%
Total	100%	100%	100%	100%
Dollar overclaim percentage[1]	30.2%	29.0%	28.1%	28.5%
	1.3%	1.2%	1.0%	0.7%

Notes: Figures may not sum due to rounding. Standard errors are presented below estimates.
1 The dollar overclaim percentage is not the same as the improper payment rate, which is calculated on an annual fiscal year basis by a different methodology and accounts for amounts that are recovered by IRS enforcement activities.

To provide a sense of how large overclaims are at the individual taxpayer level, Figure 2 shows the distribution of overclaim returns by the size of the overclaim. Figure 3, which shows the distribution of original EITC claims by size, is provided for comparison. A large percent of overclaims are less than $500: 38 percent according to the higher estimates and 44 percent according to the lower estimates.[43] This compares with 27 percent of the original claims. At the other end of the spectrum, the percent of overclaims that are greater than $3,000 is disproportionately low: 15 percent according to the higher estimates and 11 percent according to the lower estimates, compared with 23 percent of original claims.

[43] Recall that the terms "higher" and "lower" refer to the level of noncompliance rather than the numeric value of the estimate, so that here the "higher" estimate produces a lower number, indicating a lower proportion of overclaims in this low-dollar range.

Figure 2: Distribution of Overclaim Returns by Size of Overclaim
Weighted Population Estimates, Annual Averages, NRP TY 2006-2008
(Constant 2008 dollars)

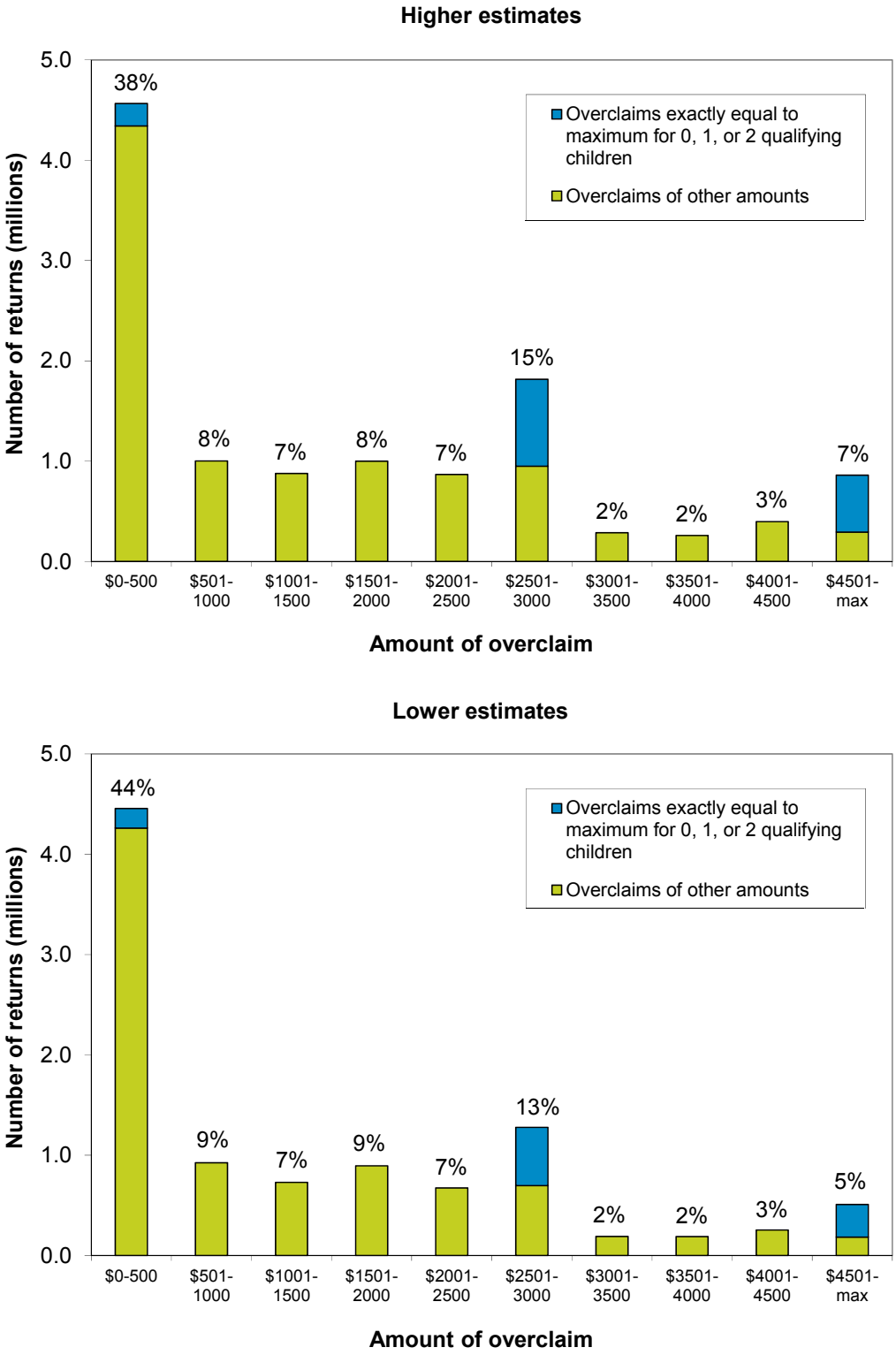
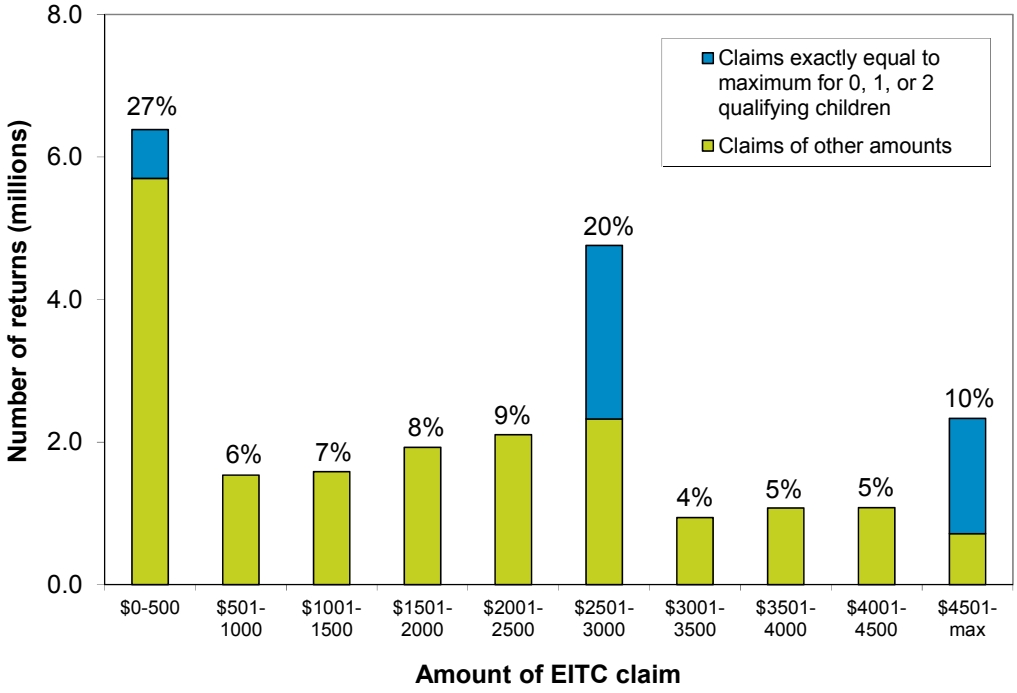


Figure 3: Distribution of EITC Claims by Size of Claim
Weighted Population Estimates, Annual Average, NRP TY 2006-2008
(Constant 2008 dollars)



Sources of errors

This section focuses on the frequency and magnitude of particular errors that are associated with the overclaim of EITC. As with the earlier 1999 Compliance Study, this analysis begins by distinguishing overclaims arising from *known* errors, which were determined during an audit in which the taxpayer fully participated, and unknown errors, for taxpayers who did not participate in the audit.[44] The category of unknown errors also includes cases where the taxpayer did not meet with the auditor or supply documentation, but eventually participated in the audit by agreeing to the changes proposed by the auditor and signing the final report. The breakdown between known and unknown errors is shown in Table 3. Note that the total dollars of overclaims described in this table ($19.3 billion) is the same figure as total overclaims in Table 1 and in Table 2a (the higher estimates). Roughly 8.4 million returns are estimated to have an overclaim with a known error, for a total of $11.4 billion in overclaims. Up to another 3.6 million returns have an overclaim with an unknown error, for another $7.9 billion in overclaims, or 41 percent of the total.

[44] Although IRS may be able to determine some errors made by the audit non-participants by using internal or third-party data, no such information is presented since it would be an incomplete and potentially skewed analysis.

Table 3. EITC Overclaims: Known and Unknown Errors
Weighted Population Estimates, Annual Average, NRP TY 2006-2008
(Constant 2008 dollars)

	Returns with EITC overclaim		Total Dollars of EITC Overclaims		Average Overclaim
	Number (millions)	Percent (%)	Dollars (billions)	Percent (%)	
Total returns with EITC overclaims	**11.9**	**100%**	**$19.3**	**100%**	**$1,614**
	0.1	--	0.4	--	$28
Type of error unknown[1]	3.6	30%	$7.9	41%	$2,214
	0.1	1%	0.3	2%	$47
Type of error known	8.4	70%	$11.4	59%	$1,360
	0.1	1%	0.3	2%	$30

Note: Figures may not sum due to rounding. Standard errors are presented below estimates.
1 Unknown errors are ones where the taxpayer did not participate in the audit or participated only by signing the final audit report.

Consider the various types of (known) errors that can lead to an overclaim of the EITC. The size of the credit is determined by earned income, AGI, number of qualifying children, and filing status; therefore, misreporting any of these items can result in claiming the wrong amount of the credit, including claiming a positive credit when the correct amount is $0. Beyond these factors, there are a number of eligibility criteria that may cause the full amount of the EITC to be disallowed, either during return processing or during an audit. These eligibility criteria will be outlined in more detail later in this section. Qualifying child errors – where a child claimed is not the taxpayer's qualifying child for purposes of EITC – seem to straddle these two categories of error in that there is both an element of eligibility and an element that contributes to the size of the credit.

The 1999 Compliance Study demonstrated that income misreporting and qualifying child errors were the two most frequent errors with the largest dollar impact on overclaims for returns filed in tax year 1999.[45] Table 4 shows that this remains true in TY 2006-2008. The Appendix provides a side-by-side comparison of the sources of errors summarized in Tables 3 and 4 with findings from the 1999 Compliance Study. For purposes of Table 4, returns with known EITC errors are split into four distinct groups: those with income misreporting, those with qualifying child errors, those with both of these errors, and those with neither. The first three of these groups are further broken down by whether additional errors are present. In this table, "income misreporting" includes the misreporting of earned income or AGI affecting the amount of the credit as well as the underreporting of investment income. Cases where the taxpayer filed as single or head-of-household, but should have filed jointly with their spouse and reported the combined income are considered "other errors" (filing status errors) rather than income misreporting, unless the income of the taxpayer or spouse was also misreported.[46] Tiebreaker errors, which result from

[45] See Table 2, page 13 in the IRS report *Compliance Estimates for Earned Income Tax Credit Claimed on 1999 Returns.*
[46] This aspect of the definition of income misreporting differs from the definition used in Table 2 of the 1999 Compliance Study, in which such cases were included with income misreporting. The number of affected cases is small: there are 37 overclaim cases in the TY 2006-2008 EITC subsample where the taxpayer filed as single or head-of-household but should have filed jointly with their spouse. Of these, 23 correctly reported income (albeit on separate tax returns) and so are treated as "other errors" in this study, while they would have been considered cases of income misreporting in the 1999 Compliance Study.

the wrong taxpayer claiming an otherwise eligible qualifying child, are considered to be "other errors" in this table rather than qualifying child errors.

Table 4. Distribution of Overclaims With Known Error By Presence of Income Misreporting[1] and Qualifying Child (QC) Errors, Weighted Population Estimates, Annual Average, NRP TY 2006-2008

(Constant 2008 dollars)

Error type	Returns with EITC overclaim		Total Dollars of EITC Overclaims		
	Number (millions)	Percent (%)	Dollars (billions)	Percent (%)	Average Overclaim
Total overclaim returns with known error	**8.4**	**100%**	**$11.4**	**100%**	**$1,360**
	0.1	--	0.3	--	$30
Income misreporting but no QC errors	**4.9**	**58%**	**$3.9**	**35%**	**$807**
	0.1	1%	0.2	1%	$31
Income misreporting alone[2]	4.3	51%	$2.9	25%	$673
	0.1	1%	0.1	1%	$30
In combination with other errors[3]	0.6	7%	$1.1	9%	$1,737
	0.0	1%	0.1	1%	$93
QC error(s) but no income misreporting	**1.8**	**21%**	**$4.3**	**38%**	**$2,384**
	0.1	1%	0.2	2%	$50
Qualifying child error(s) alone[4]	1.3	15%	$3.0	26%	$2,327
	0.1	1%	0.2	2%	$68
In combination with other errors	0.5	6%	$1.3	11%	$2,529
	0.0	1%	0.1	1%	$98
Both income misreporting and QC error(s)	**0.7**	**9%**	**$1.7**	**15%**	**$2,451**
	0.1	1%	0.2	1%	$114
Income and qualifying child error(s) only	0.5	6%	$1.3	12%	$2,513
	0.0	1%	0.1	1%	$136
In combination with other errors	0.2	2%	$0.4	4%	$2,275
	0.0	0%	0.1	1%	$231
All other errors (no income or QC errors)	**1.0**	**12%**	**$1.4**	**12%**	**$1,447**
	0.1	1%	0.1	1%	$86

Note: Figures may not sum due to rounding. Standard errors are presented below estimates.

1 Income misreporting includes both misreporting of the amounts used to calculate the credit (i.e., earned income or AGI) as well as underreporting of investment income.

2 This category includes returns where the number of qualifying children was actually increased during audit, so to some extent the underclaim of the qualifying children offsets the effect of income misreporting; however these are all net overclaim cases.

3 For purposes of this table, "other errors" includes the following: filing status errors, errors corrected in processing, tiebreaker errors, and one of these violations of eligibility criteria: invalid SSN for taxpayer, lack of U.S. citizenship or resident alien status for the full year, filing of Form 2555 or 2555-EZ, and these errors specifically for taxpayers claiming EITC with no children: age other than 25-64, being the dependent or qualifying child of another person. "Other errors" also includes claiming the credit while a ban was in place and claiming the credit without recertifying if EITC was denied in a previous year. IRS Publication 596 details all EITC eligibility criteria.

4 This category includes returns where there was some income misreporting but it was not in the taxpayer's favor and it was not enough to offset the effect of claiming children that were not eligible. This category also includes some operational exam cases where there is limited information about the errors that occurred; what is known is that these cases had EITC fully disallowed and there was no audit adjustment to income or filing status, so the qualifying child error (only) is assumed.

As Table 4 indicates, income misreporting is the most commonly made error, occurring on about 67 percent of overclaim returns with a known error.[47] In most of those cases – roughly half of overclaim returns with a known error – income misreporting is the *only* error. Overclaim dollars associated with income misreporting (only) are disproportionately much lower, at 25 percent. This is to be expected, since income misreporting can often cause comparatively small changes to the amount of the credit, for an average overclaim of $673, relative to eligibility errors in which the full credit is disallowed, which average upwards of $2,000. Qualifying child errors (excluding those that occur alongside income misreporting) show this pattern in reverse: these errors represent a much higher percent of overclaim dollars (38 percent) than overclaim returns (21 percent). Where the only error is a qualifying child error, the average overclaim is $2,327.

The chances are relatively low that a return with an error will have neither income misreporting nor a qualifying child error – just 12 percent. There is also a fairly low chance that a return will have both types of errors; this happens on only 9 percent of returns with known errors. Although "other errors" – those aside from income misreporting and qualifying child errors – have taken a backseat in this discussion so far, they are nontrivial as a group: adding together the relevant rows in Table 4 (those with "other errors" in the description) indicates that these appear on 27 percent of returns, and these returns account for 37 percent of overclaim dollars.

The high rate of errors on EITC returns is often explained by the complexity of the credit, particularly with respect to how it applies when family structures and living arrangements do not conform to the traditional nuclear family.[48] While Table 4 cannot fully speak to the role of complexity in contributing to noncompliance, it does show that about 50 percent of the errors and at least 25 percent of the dollar overclaims stem from income misreporting alone and therefore cannot reasonably be attributed to the complexity of family living situations. This does not imply that income reporting for EITC is free from complexity, although one would expect this to be much less true since TY 2002, when the modified AGI concept was abandoned and the simplified definition of earned income was introduced.

In fact, the TY 2006-2008 NRP data provide good reason to believe that the income misreporting errors reported in Table 4 generally do not stem from complexity created by the EITC. This is because, for most taxpayers, the computations required to determine income relevant for the credit (earned income, AGI, investment income) appear to be fairly straightforward, if one excludes the elements of those computations that must be made regardless of whether EITC is claimed. For instance, earned income must be calculated for purposes of EITC, but for 96 percent of EITC claimants, the correct amount of earned income can be determined by combining at most 4 line items from the front of the Form 1040.[49] In fact, for roughly three-

[47] The 67 percent is arrived at by summing the row labeled "Income misreporting but no QC errors" (58 percent) and the row labeled "Both income misreporting and QC error(s)" (9 percent).

[48] The National Taxpayer Advocate states that "most EITC errors result from applying the complexity of the EITC rules to the complexity of families' lives," (National Taxpayer Advocate, 2011, pg 303); Holt (2006) observes that "Potential EITC claimants face many … problems in interpreting what are often complicated lives through the lens of the tax code," (page 19). The complexity of the EITC has also been ascribed not so much to the credit by itself but to its interaction with other child-related tax-benefits, which together form a patchwork of eligibility criteria that can be challenging to navigate (Holt, 2006; Holtzblatt and McCubbin, 2003; Maag 2011). Certainly this difficulty would be exacerbated by complicated family situations.

[49] The four line items are wages, business income, farm income (or loss), and one-half the self-employment tax. In TY 2008, these correspond to lines 7, 12, 18, and 27 on the Form 1040. To obtain the 96 percent figure in the main text, earned income was calculated from these four lines (described below) and compared with the "Earned Income"

quarters of EITC claimants, earned income is identical to the amount reported on the wages line. Similarly, for 97 percent of taxpayers claiming EITC, investment income can be calculated by summing 4 lines on the front of the Form 1040.[50] Thus, for most taxpayers, the additional steps required to calculate income concepts for the EITC are fairly simple.[51]

From the standpoint of administering the credit, it is important to understand whether the income misreporting that contributes to EITC overclaims is caused by the EITC itself (due to complexity or otherwise) or whether it is the kind of income misreporting that occurs across a wider population of taxpayers. In particular, it seems worth exploring the extent to which the credit itself may be causing taxpayers to manipulate their reported income to be eligible for (a larger amount of) the credit.[52] This has implications for whether these errors should be addressed within the context of EITC or whether broader resource allocation is needed to resolve these errors.

While the overview of errors in Table 4 shows the relative importance of income misreporting and qualifying child errors, the next set of results provides information for a wider set of errors and better informs how much each type of error by itself contributes to total overclaimed dollars. Attributing overclaims to error types is straightforward for returns that contain only one error, but can be challenging for returns that have multiple errors – and, as Table 4 indicates, multiple errors are common.

Take as an example a case where the taxpayer files as head-of-household and claims $2,000 in EITC, but whose correct status is married-filing-separately and who has investment income above the threshold. If, hypothetically, one of these errors could be prevented or recovered – if say, the taxpayer were deterred from incorrectly reporting filing status or if the IRS detected the error – the full $2,000 overclaim would be prevented or recovered. This same outcome would occur if instead the investment income error were prevented or detected. This suggests it might

---

value that is computer-generated by IRS during return submission processing. When these figures do not match, it implies other line items contribute to the determination of earned income. In only 4 percent of cases, the figures do not match, implying that only 4 percent of the EITC population faced additional complexity in calculating earned income. Earned income was calculated as follows: for those not reporting any self-employment income, it equaled the amount on the wages line; for those reporting self-employment income but no wages, it equaled the sum of business income and farm income minus one-half the self-employment tax; for those with both wages and self-employment income, the two previous calculations were summed together.

[50] Only 3 percent of the EITC-claiming population file one of the forms or schedules that might complicate the calculation of investment income (Schedule E, Form 4797, or Form 8814). The four lines on the Form 1040 that comprise investment income for the other 97 percent are taxable interest, tax-exempt interest, ordinary dividends and any positive capital gains (losses are zeroed out); these are reported on lines 8a, 8b, 9a and 13 on the 2008 Form 1040.

[51] This conclusion is despite the challenges that taxpayers may face in determining the proper amount of each contributing line item, such as business or farm income. But because the items underlying earned income as well as investment income and AGI would need to be calculated and reported even if the EITC were not claimed on the return, it seems inappropriate to attribute any complexity associated with determining these amounts to the EITC.

It should also be noted that, even if the computations end up being straightforward, the taxpayer may still have to work through instructions of more complex situations in order to ensure that the simple calculations are appropriate.

[52] Recent work by Chetty, Friedman, Ganong, Leibel, Plumley and Saez (2011) shows that much of the observed "bunching" of self-employed taxpayers (i.e., reporting income in the range that maximizes the credit and minimizes tax liability) is largely due to income misreporting rather than labor supply effects. This shows that enough taxpayers are targeting their income misreporting to the specific parameters of the EITC to have measurable effects. More work would need to be done to try to assess what share of income-misreporting overclaims could be attributed to this kind of behavior.

be reasonable to attribute $2,000 of overclaims to filing status error and $2,000 to investment income error. The drawback with this approach is that it implies that total overclaims are $4,000. Alternatively, one might consider the hypothetical effects of preventing or detecting the errors sequentially and attributing dollars of prevented or recovered overclaims accordingly: if one considered preventing the filing status error first and investment income second (ignoring all other error types for the sake of this example), one would attribute all $2,000 to filing status error, leaving $0 to be attributed to investment income error. Yet not only does this approach beg the question of the appropriate sequence in which to consider the various errors, it misleadingly suggests, in this example, that investment income error is not responsible for any amount of overclaims.

In the absence of one simple way to assign overclaims to error types, the approach taken in the next set of tables is to provide *two* estimates for each error type: the first shows how many dollars of overclaims would be prevented or recovered if, hypothetically, the IRS were to deter or detect every instance of that error. The second shows how many dollars of overclaims would be prevented or recovered if the given error were deterred or detected only after all other errors were deterred or detected. (For returns with only one error, the first and second estimates will be the same.) Returning to the earlier example, the first estimate for filing status error would be $2,000 and the second estimate would be $0. Likewise, the first and second estimates for investment income error would also be $2,000 and $0. The two estimates can be considered roughly the maximum and minimum dollar amounts that would be prevented or recovered if hypothetically (and implausibly) the IRS were able to eliminate all instances of that type of error, either through deterrence or detection.[53]

Table 5a provides these estimates along with the frequency of each category of error, shown in the first column.[54] Table 5b presents a similar analysis to Table 5a, but instead of attributing *total* overclaims across the population to each error type, *average* overclaims per error are presented instead. As shown, the error types are broken out into more detail than they were in Table 4. In this table, income misreporting errors are separated by whether the misreported amount involves earned income or AGI or investment income.[55] Earned income misreporting is

[53] There are some unusual or unexpected situations that mean this is not a hard-and-fast rule. For instance, in some cases preventing or detecting income misreporting actually increases the EITC; while doing so after preventing or detecting incorrect claims of the number of qualifying children reduces the EITC or has no effect. In that case, the first "maximum" estimate would be smaller (negative) than the second "minimum" estimate (positive or zero). Given that this analysis is limited to returns with a net overclaim, this type of situation does not occur with much frequency. The more common exception to the above statement occurs for eligibility criteria for EITC without children. This is because, for taxpayers that had qualifying children disallowed during audit, breaking a rule for EITC without children would not have any impact if it were the only error deterred or detected. But if an incorrect claim of number of qualifying children were prevented or detected first, then breaking a rule for EITC without children would wipe out the remaining credit. This type of error is common enough to make the "minimum" estimate actually exceed the "maximum" estimate.

[54] This column only counts those errors that actually result in an overclaim on net overclaim returns. An example of a fairly common error not included here is when a taxpayer files as head-of-household but should have filed as single; since this particular filing status error does not affect the amount of the credit due the taxpayer, it is not counted in this column. Likewise, some earned income errors do not affect the amount of the credit if the reported and corrected amounts are both within a certain range. Also excluded are errors that appear on returns that do not have a net overclaim; for example, if a processing (math) error reduces the credit but this is offset by another correction during audit that yields a net increase to the EITC, that processing error will not be included here.

[55] The misreporting of other income sources can affect the amount of the EITC in two ways: first, there are maximum thresholds for AGI and investment income above which the taxpayer is ineligible, irrespective of earned

further separated into whether the source is self-employment income or wage income, to reflect expected differences in misreporting by income type.[56] The "other errors" from Table 4 are separated into five categories: filing status errors, tiebreaker errors, errors corrected during processing, and two categories of additional eligibility criteria. The first of these two latter categories is a combined set of the remaining rules for all taxpayers (having a valid Social Security Number, being a U.S. citizen or resident alien all year, not filing Form 2555 or Form 2555-EZ to exclude foreign earned income, not being a qualifying child of another person); the second category is the set of rules for taxpayers not claiming qualifying children (being age 25 to 64 years, not being a dependent of another taxpayer, having a home in the U.S. for more than half the year).[57] Tables 5a and 5b do not separate qualifying child errors into more specific error types; this is done in a subsequent table.

Unlike Table 4, which summarizes only the known errors ($11.4 billion in overclaims), Tables 5a and 5b incorporate the unknown errors as well: an additional $2.6 billion for the lower estimates and an additional $7.9 billion for the higher estimates.[58] The additional errors and overclaim dollars are attributed to the audit non-participants based on the patterns of errors made by audit participants within the same stratum. This is done for both the higher and lower estimates.

Table 5a confirms what was already demonstrated in Table 4: the biggest contributors to overclaims are income misreporting and qualifying child errors. The relative importance of qualifying child errors in dollar terms seems even more striking. If all qualifying child errors, and *only* qualifying child errors, were prevented or detected, an estimated $10.4 billion (higher estimate) or $7.2 billion (lower estimate) in overclaims would be prevented or recovered. The

---

income. Second, in some cases the calculation of the credit amount is actually based on AGI rather than earned income, so changes to AGI that are below the threshold can still affect the amount of the credit. Note that misreported amounts of AGI that correspond to misreported earned income are excluded from the category of "other types of income misreporting." Likewise, when misreported AGI corresponds dollar-for-dollar to misreported investment income, it will be treated as a single error.

[56] Chetty, Friedman, Ganong, Leibel, Plumley and Saez (2011) demonstrate that, among taxpayers claiming EITC, income misreporting is more prevalent by those who report some self-employment income compared with those who report only wage income. More generally, income misreporting is known to be strongly associated with the extent to which income information is reported to the IRS by a third party. Self-employment income has virtually no information reporting while wage income is subject to heavy information reporting as well as withholding. Tax gap research conducted by IRS Office of Research has demonstrated the link between third-party information reporting and reporting noncompliance. For example, according to the Individual Income Tax Underreporting Gap Estimates for Tax Year 2001, the net misreporting percentage (NMP) for wages is 1.2 percent and for self-employment (nonfarm proprietor) income, the NMP is 57.1 percent (http://www.irs.gov/pub/irs-utl/tax_gap_update_070212.pdf). The NMP is the amount of income misreported divided by the amount that should have been reported.

[57] These eligibility criteria are described thoroughly in IRS Publication 596. The two sets of eligibility criteria largely follow these two chapters in that publication: Chapter 1, "Rules for Everyone," and Chapter 3, "Rules If You Do Not Have a Qualifying Child." There are two exceptions to this: first, Chapter 1 includes rules involving the thresholds of earned income and AGI, but, for purposes of this analysis, those are considered with income misreporting; second, the rule that the taxpayer cannot be the qualifying child of another person is included separately in both Chapter 3 and in Chapter 2, "Rules If You Have a Qualifying Child." Here, that criterion is considered along with the rules for all taxpayers. Note that IRS Publication 596 is published annually, but the 2006, 2007 and 2008 versions are essentially the same, aside from dollar amounts that are adjusted annually.

[58] Adding these additional overclaim dollars produces the lower and higher estimates of total overclaim dollars: $11.4 plus $2.6 and $7.9 equals $14.0 and $19.3, respectively, the totals shown in Table 1 and again in the bottom row of Table 5a here.

second set of estimates indicates that even if all other types of errors were eliminated first, there would still remain $8.5 billion (higher estimate) or $5.9 billion (lower estimate) in overclaims due only to qualifying child errors. Taking these estimates together, we find that qualifying child errors account for 42 to 54 percent of total overclaims. Table A3 in the Appendix provides similar percentages for all error types; some of these are also referred to in the discussion below.

Table 5a. Total Dollars of EITC Overclaims Attributable to Common Types of EITC-Related Errors
Weighted Population Estimates, Annual Average, NRP TY 2006-2008
(Constant 2008 dollars)

Error type	Number of returns with error[1] (millions)	Total overclaims preventable or recoverable (billions)			
		Higher estimates		*Lower estimates*	
		If this is the only error deterred or detected[2]	After all other errors are deterred or detected[3]	If this is the only error deterred or detected[2]	After all other errors are deterred or detected[3]
Error corrected in processing[4]	0.5	$0.5	$0.5	$0.4	$0.4
	0.0	0.1	0.1	0.1	0.1
All income misreporting[5]	6.5	$5.6	$4.7	$4.5	$3.8
	0.1	0.2	0.2	0.2	0.2
Earned income misreporting	4.5	$4.5	$3.5	$3.7	$2.9
	0.1	0.2	0.2	0.2	0.1
Wage income	1.7	$1.1	$0.6	$0.8	$0.4
	0.1	0.1	0.1	0.1	0.0
Self-employment income	3.1	$3.8	$2.9	$3.2	$2.5
	0.1	0.2	0.2	0.2	0.1
Other types of income misreporting	3.1	$1.5	$0.9	$1.1	$0.7
	0.1	0.1	0.1	0.1	0.1
Qualifying child error	3.0	$10.4	$8.5	$7.2	$5.9
	0.1	0.4	0.3	0.3	0.3
Tiebreaker error	0.1	$0.3	$0.2	$0.2	$0.2
	0.0	0.1	0.1	0.0	0.0
Filing status error	1.0	$3.3	$1.9	$2.3	$1.3
	0.1	0.3	0.2	0.2	0.2
Rules for all taxpayers claiming EITC[6]	0.3	$1.0	$0.1	$0.7	$0.1
	0.0	0.2	0.0	0.1	0.0
Rules for EITC without children[7]	0.3	$0.1	$0.2	$0.0[8]	$0.1
	0.0	0.0	0.0	0.0	0.0
Total overclaims	11.9[9]	$19.3		$14.0	

Note: Standard errors are presented below estimates. See Table 5b for explanation of notes 1-7 in this table.
8 Less than $50 million.
9 This figure is the higher estimate of the number of returns with at least one error leading to an overclaim, which can also be seen in Table 1. The comparable lower figure is 10.1 million.

Table 5b. Average Dollars of EITC Overclaims Attributable to Common Types of EITC-Related Errors
Weighted Population Estimates, Annual Average, NRP TY 2006-2008
(Dollars in constant 2008 dollars)

Error type	Number of returns with error[1] (millions)	Average overclaim dollars[8]			
		Higher estimates		*Lower estimates*	
		If this is the only error deterred or detected[2]	After all other errors are deterred or detected[3]	If this is the only error deterred or detected[2]	After all other errors are deterred or detected[3]
Error corrected in processing[4]	0.5	$1,089	$1,089	$783	$783
	0.0	154	154	103	103
All income misreporting[5]	6.5	$860	$716	$690	$581
	0.1	34	31	28	25
Earned income misreporting	4.5	$1,018	$780	$831	$648
	0.1	45	35	37	30
Wage income	1.7	$671	$365	$480	$258
	0.1	69	36	50	25
Self-employment income	3.1	$1,237	$946	$1,040	$809
	0.1	55	45	46	39
Other types of income misreporting	3.1	$488	$306	$368	$230
	0.1	35	25	26	18
Qualifying child error	3.0	$3,424	$2,812	$2,388	$1,952
	0.1	64	76	41	49
Tiebreaker error	0.1	$3,375	$2,825	$2,323	$1,945
	0.0	424	444	279	292
Filing status error	1.0	$3,252	$1,873	$2,264	$1,297
	0.1	144	163	96	109
Rules for all taxpayers claiming EITC[6]	0.3	$2,866	$437	$2,008	$304
	0.0	270	133	182	91
Rules for EITC without children[7]	0.3	$183	$552	$130	$386
	0.0	23	124	16	84
Average overclaim per overclaim return	--	$1,614		$1,390	

Note: Standard errors are presented below estimates. See Table 5a for explanation of notes 1-7 in this table.
1 These are the estimates of errors that contribute to overclaims on net overclaim returns only.
2 In general, this is the maximum attribution of dollars to each error type. See Footnote 53 for discussion of exceptions to this. The column total (not presented) would exceed the estimate of total overclaim dollars, shown in final row.
3 In general, this is the minimum attribution of dollars to each error type. See Footnote 53 for discussion of exceptions to this. The column total (not presented) would fall short of the estimate of total overclaim dollars, shown in the final row.
4 This category primarily consists of math error but includes other adjustments made before the NRP audit. Since these errors are detected before any of the others, there is no logical problem here with counting them independently of sequence.
5 The values for "all income misreporting" do not equal the sum of the values for wage income, self-employment income, and AGI and investment income for the same reason one would not expect the column totals for the full table to match the estimates for the population provided in the final row. Similarly, the values for earned income misreporting do not equal the sum of the values for wage income and self-employment income.
6 This category consists of eligibility rules not previously listed that apply to all taxpayers regardless if they are claiming children. It includes having a valid SSN, being a U.S. citizen or resident alien all year, not filing Form 2555 or Form 2555-EZ, and not being a qualifying child of another person. These rules are outlined in detail in IRS Publication 596.
7 This category consists of eligibility rules that apply to taxpayers claiming EITC without children. It includes being age 25-64, not a dependent of another taxpayer, and having a home in the U.S. for more than half the year. These are outlined in detail in IRS Publication 596.
8 Overclaims are averaged only over the returns that have the type of error in question.

Income misreporting is the second biggest contributor to overclaims, accounting for 24 to 32 percent of total overclaims. Among income types, self-employment income misreporting is the most significant contributor to overclaims (15 to 23 percent), with wage income misreporting being the least significant (3 to 6 percent). This is the reverse of how these sources of earned income are represented in the EITC-claiming population, where wage income is far more common: 76 percent of taxpayers claiming EITC earn only wage income, while the remaining 24 percent earn at least some self-employment income (10 percent report both wages and self-employment income).[59]

The relative importance of self-employment income contrasts with existing published research, which has indicated that the misreporting of self-employment income, strategic or otherwise, does not play a significant role in EITC noncompliance (McCubbin, 2000; Hotz and Scholz, 2003). Yet the existing research has largely relied on outcomes from one of the first EITC compliance studies, *Study of EITC Filers for Tax Year 1994* (IRS, 1997). There are several possible reasons why the TY 2006-2008 NRP data produce different conclusions. First, the share of EITC claimants reporting self-employment income has risen since 1994, from roughly 15 percent to 24 percent in 2006-2008, which by itself would be expected to increase the relative importance of self-employment income misreporting. Second, the 1994 study may have understated its importance in the first place. The sample for the 1994 study substantially underrepresented taxpayers reporting self-employment income, even after sample weighting: while Statistics of Income (SOI) data showed that 15 percent of EITC claimants reported Schedule C income in TY 1994, only 6 percent of the sample in the 1994 Compliance Study reported some Schedule C income.[60]

The third possible reason for the difference is that the data for the 1994 study were collected by agents of the IRS Criminal Investigation Division (now referred to as CI) rather than tax auditors. CI agents typically detect instances of fraud and other forms of noncompliance that are considered more serious than income misreporting.[61] For the 1994 Compliance Study, they gathered information primarily through interviews with taxpayers and their associates (e.g., employers, return preparers, family members, and neighbors), which may have been more likely to uncover instances of EITC ineligibility rather than misstatements of income. In comparison, the NRP – whose main purpose is to detect unreported income through audits of books and records – could potentially provide more reliable data on income misreporting.

Filing status errors emerge as the third-largest contributor to overclaims, accounting for between $1.3 billion and $3.3 billion dollars in overclaims (9 to 17 percent of overclaims) and falling somewhere between self-employment income misreporting and other types of income misreporting in relative importance. Most of these overclaims come from married taxpayers who file separately from their spouse and incorrectly claim either single or, more frequently, head-of-

[59] For these percentages, the type of earned income is based on the correct type of income determined during audit, not what was reported on the original tax return, although the differences are slight: reported wage-only earners are 77 percent of the EITC population, with 12 percent reporting both wages and self-employment income.
[60] See 1994 Compliance Study, Table 1, page 9 for this comparison. The underrepresentation of Schedule C taxpayers in the 1994 Compliance Study sample may have to do with its sampling time frame, which only lasted through April. This may not have allowed enough time to pass after the April 15 filing deadline to process all Schedule C returns.
[61] Within the IRS, "fraud" is more than a general term used to describe income misreporting; it applies to noncompliance that has risen to a certain level of egregiousness.

household filing status. This practice tends to overstate the amount of the credit on one or both returns by splitting household income. Five percent of all EITC claimants (2 percent of those filing single and 9 percent of those filing as head-of-household) are estimated to have the correct status of married-filing-separately, making them ineligible for the credit.[62]

One may also observe from Table 5a that tiebreaker errors appear negligible in comparison with other error types (1 to 2 percent of all overclaims). This contrasts with the findings of the 1999 Compliance Study, in which tiebreaker errors were shown to be one of the most common errors, accounting for 17 percent of overclaims ($1.6 billion, 1999 Compliance Study, Table 2, page 13). This difference reflects the effects of EGTRRA, effective in TY 2002, which simplified the tiebreaker rules and effectively legislated a form of noncompliance out of existence (or into compliance). As previously noted, prior to 2002, the law dictated that when more than one person had the same qualifying child, only the one with the highest modified AGI was entitled to claim the child; beginning in 2002, taxpayers could decide amongst themselves who would claim the child. This means the type of error where the "wrong" taxpayer claims a child for EITC was eliminated. The remaining type of tiebreaker error reflected in the NRP TY 2006-2008 sample and in Tables 5a and 5b is when more than one person actually claims the child, so that at least one taxpayer must have claimed the child in error.[63]

Some of the gains toward simplification made by EGTRRA through modification of the tiebreaker rules have been mitigated by new legislation effective in TY 2009 that removes taxpayers' discretion to decide who claims a child when more than one person has the same qualifying child. This was a provision of the *Fostering Connections to Success and Increasing Adoptions Act of 2008*, entitled "Clarification of Uniform Definition of a Qualifying Child" (Joint Committee on Taxation, 2009). With the new provision in place, the law essentially restricts qualifying child tax benefits to the parents of the child, with the exception that if the parents may claim the child but do not do so, another individual can claim the child only if that individual has a higher AGI than any parent and/or any other individual who is also eligible to claim the child.

Prior to TY 2009, these same rules existed but were applied only if more than one taxpayer actually claimed the child; this meant that the complicated tiebreaker rules could be ignored by taxpayers who agreed between themselves who would claim the child. As of TY 2009, tiebreaker rules cannot be ignored whenever more than one individual has the same qualifying child; instead they have to be read and understood by all of these taxpayers in order for them to ensure they are complying with the law.

The intent behind this change may have been to limit the extent to which taxpayers could strategically claim a qualifying child in order to obtain the highest possible credit, or it may have

[62] Roughly one-fifth of these, or 1 percent of the total EITC sample, chose to change their filing status to married-filing-jointly as part of the resolution of the audit rather than maintain two married-filing-separately returns. In some cases this made the taxpayers eligible for a smaller amount of the EITC than was originally claimed, rather than fully ineligible.

[63] The 1999 Compliance Study points out that it does not account for offsetting errors when more than one taxpayer resided with the child but the wrong person (i.e., the one with lower modified AGI) claimed the child. Given the legislative change allowing taxpayers to choose who claims the child, this type of offsetting error no longer exists. However, there may remain situations where a child that was claimed incorrectly for the EITC for reasons other than the tiebreaker rules could have been claimed correctly by another taxpayer.

been to elevate the rights of the parent. According to the Joint Committee on Taxation, the new provision was estimated to save just under $200 million in revenue per year through FY 2018 (Joint Committee on Taxation, 2009). Yet another expected result of this change in legislation is a rise in tiebreaker errors, which may be reflected in future studies of EITC compliance. Even if some or most taxpayers comply with the new law, many taxpayers will continue the same behavior that was considered compliant before the legislative change, either knowingly or unknowingly contravening the new law. Those new to the EITC will find the tiebreaker rules substantially more complicated and will likely make errors as a result.

This illustrates the tradeoff between revenue savings that result from legislation geared toward preventing strategic or erroneous EITC claims and the associated costs of such legislation, including increased tax compliance burden faced by taxpayers, a rise in improper payments, and increased burden on the IRS due to the difficulty of administering certain rules.

Qualifying child errors

This section presents additional detail about the nature of qualifying child errors. The first table, Table 6, provides a summary of outcomes at the tax-return level (as opposed to the qualifying-child level). As shown, at least 70 percent of returns claiming EITC with qualifying children claimed the correct number of children, with up to another 15 percent (the audit non-participants) possibly claiming the correct number. This translates into between 13 and 27 percent of children being claimed in error, shown in Table 7.[64]

Table 6. Summary of Outcomes for Returns Claiming Qualifying Children (QC)
Weighted Population Estimates, Annual Average, NRP TY 2006-2008

	Number of returns (millions)	Percent of returns claiming QC	Number of QC (millions)	Percent of QC initially claimed
Total returns claiming QC	18.6	100%	28.2	100%
	0.0	--	0.0	--
Returns represented by audit non-participants	2.8	15%	4.1	15%
	0.1	1%	0.2	1%
Returns represented by audit participants	15.8	85%	24.0	85%
	0.1	1%	0.2	1%
Who claimed correct number of QC	13.0	70%	19.8	70%
	0.1	1%	0.2	1%
Who claimed at least 1 QC in error	2.8	15%	4.2	15%
	0.1	1%	0.2	1%
Number of QC correctly claimed on those returns			0.4	2%
			0.0	0%
Number of QC claimed in error			3.8	13%
			0.1	1%

Notes: Figures may not sum due to rounding. Standard errors are presented below estimates.

[64] This excludes the few children who were claimed in error, but where another child in the family was established as a qualifying child during the audit, meaning the taxpayer claimed the correct *number* of qualifying children, if not the right children themselves.

Table 7 explores the nature of the qualifying child errors. During this period, there were three "tests" for qualifying child eligibility: the relationship, age, and residency tests, all of which are still applicable in 2013. To meet the relationship test, the child has to be the taxpayer's son/daughter, niece/nephew, sibling, foster child, or a descendant of any of these. In TY 2006-2008, the relationship test also required the child to be unmarried, unless certain special conditions were met; for purposes of this analysis, "married child" is treated as a separate error.[65] To meet the age test, the child has to be either younger than 19, younger than 24 and a full-time student, or any age and permanently and totally disabled.[66] To meet the residency test, the child has to live with the taxpayer in the U.S. for more than half the year. Although not considered part of the definition of a qualifying child, in order to be eligible for EITC, the child must also have a valid Social Security Number (SSN). Although tiebreaker errors are not technically qualifying child errors, they are included in this table.[67] These eligibility criteria, including special exceptions or qualifications, are described in greater detail in IRS Publication 596, Chapter 2: Rules If You Have a Qualifying Child.

Two additional types of error are included in this analysis: "errors corrected during processing" and unknown errors. The first group consists primarily of math errors that are specific to qualifying children: these children can be disallowed during return processing if third-party information suggests that the age requirement is not met or the child's SSN is not valid, or if information reported on the Schedule EIC suggests the child is not eligible for the credit. However, if the math error was reinstated before the start of the NRP audit based on additional information provided by the taxpayer, it will not be included here unless the case has other errors yielding a net EITC overclaim at the end of the audit. The "errors corrected during processing" category also includes any other changes to the EITC amount made between the initial claim on the filed return and the start of the NRP audit, if that change demonstrated a qualifying child error, such as when the taxpayer filed an amended return and removed the child from the return.

The unknown errors are reported separately as either "not substantiated" or "unknown error." Neither of these are the same as errors that come from audit non-participants, reported separately in this analysis. For these unknown errors, there is good reason to believe that a qualifying child error was made, but there is no further information about the nature of the error. The "not substantiated" errors reflect cases where the taxpayer did not meet with the auditor or supply documentation, but ultimately agreed to and signed the auditor's report disallowing the children. Thus, the taxpayer acknowledged that the children were not eligible, but it cannot be determined which eligibility criteria were violated. The second "unknown" category is primarily made up of returns that were audited by operational exam, so no detail was collected about the specific errors.[68] It also includes a handful of NRP audits where the audit is not documented well enough to determine which error was made.

---

[65] Beginning in TY 2009, the requirement that the child not be married was instead incorporated into a fourth test, the joint return test.

[66] Beginning in TY 2009, the age test added the requirement that the child also has to be younger than the taxpayer or taxpayer's spouse.

[67] By definition, a tiebreaker error can only occur if the claimed child meets the criteria to be a qualifying child of more than one taxpayer.

[68] Recall that NRP sample selection occurs after other steps in the return processing pipeline, and in particular after returns are selected for pre-refund audit. These operational exam cases are audited according to standard procedures and are not subject to the additional data collection requirements of the NRP. The operational exam cases that are

Table 7 presents the estimated population frequency of each of these qualifying child error types. This table does not indicate when errors overlap, although it can be seen that substantial overlap exists. Of the known errors, the largest error is a failure to meet the residency test. At least 75 percent of the children known to be claimed in error fail the residency test; this is roughly 10

Table 7. Frequency of Specific Qualifying Child (QC) Errors
Weighted Population Estimates, Annual Average, NRP TY 2006-2008

	Number of qualifying children (millions)	Percent of all QC claimed	Percent of all QC claimed in error by audit participants
Total QC claimed	28.2	100%	
QC claimed (possibly in error) by audit non-participants	4.1	15%	
QC claimed in error by audit participants	3.8	13%	
Frequency of errors estimated from audit participants			
Relationship	0.8	3%	20%
	0.1	0%	2%
Age	0.4	1%	10%
	0.0	0%	1%
Residency	2.8	10%	75%
	0.1	0%	4%
Invalid SSN	0.3	1%	8%
	0.0	0%	1%
Married child	0.0[3]	0%	1%
	0.0	0%	0%
Tiebreaker	0.3	1%	7%
	0.0	0%	1%
Error in processing	0.0[3]	0%	1%
	0.0	0%	0%
Not substantiated[1]	0.2	1%	4%
	0.0	0%	1%
Unknown error(s)[2]	0.3	1%	7%
	0.0	0%	1%

Notes: Standard errors are presented below estimates. Columns will not sum to total given occurrence of multiple errors on returns.
1 This category of errors consists of those where the taxpayer does not initially respond to communication from the examiner, but ultimately signs and agrees to the examiner's final report that disallowed the child(ren).
2 This category of errors includes cases worked by standard operational audit (non-NRP) where no change to income or filing status is observed; a qualifying child error is presumed, but no detail is available. This category also includes cases where the specific error made cannot be determined from typical data collection instruments for NRP or from the electronic forms of audit documentation.
3 Less than 50 thousand.

assumed to be qualifying child errors fall into one of two camps: the larger of the two consists of audits where the full amount of the EITC was disallowed, but no income error or filing status error is observed. It may therefore be somewhat overstated because there may be other eligibility criteria violated aside from the eligibility of the children. However, even the number of cases involved in this "larger camp" is small. The smaller camp consists of cases where the final EITC amount is positive and consistent with the amount of the credit that would be due the taxpayer with the audit-corrected level of income and no qualifying children.

percent of all children claimed for EITC. The relationship test is the second most frequent error: of those children known to be claimed in error, 20 percent fail to meet the relationship test, which is roughly 3 percent of all children claimed. The remaining errors each affect 10 percent or less of children claimed with known error, or roughly 1 percent or less of all children claimed. Additional work can be done to delve more deeply into these less common error types, and the NRP 1040 Study provides opportunity for such analysis.

Tax return preparers

One area of interest to the IRS is the relationship between tax return preparation services and EITC errors made on filed returns. The NRP study collects more detail on preparer type than is typically available, which provides an opportunity to begin comparing EITC errors across preparer types.

Table 8 shows how frequently EITC claimants use different preparer types, compared with preparer usage for returns not claiming EITC. There is a sizable difference in the tendency to self-prepare the return, with 43 percent of non-claimants preparing their own returns and 29 percent of EITC claimants self-preparing.[69] Among those who reported using a particular type of paid preparer, shown in the third and sixth columns of Table 8, EITC claimants are more likely to use an unenrolled return preparer (43 percent) or a preparer from a national tax return preparation firm (35 percent) than non-claimants (28 percent and 14 percent, respectively). In contrast, non-claimants for EITC are much more likely to use a CPA to prepare their return: 44 percent do so. This compares with just 10 percent of EITC claimants.

Table 9 shows how EITC errors differ across preparer types. The first three rows compare outcomes between self-prepared returns, paid-preparer returns, and returns prepared by the IRS or IRS-sponsored programs.[70] There is no statistical difference between self-prepared and paid-preparer returns in either the frequency of overclaims or the dollar overclaim percentage. Returns prepared by the IRS or its sponsored programs, Volunteer Income Tax Assistance (VITA) and Tax Counseling for the Elderly (TCE), have much lower frequency of error and overclaim percentages than other preparer types. These provide an interesting benchmark for considering the performance of other preparer types. Unlike taxpayers and paid preparers, the volunteers in these programs have no incentive to overstate EITC claims; arguably they could be more familiar with current tax law than others, having been specially trained by the IRS. It may not be surprising, then, that these returns appear to have the lowest errors, whether measured by percent of returns with an overclaim (26 percent higher estimate, 20 percent lower estimate) or the dollar overclaim percentage (13 percent higher estimate, 11 percent lower estimate).

[69] According to more recent data, the rate of self-preparation among EITC claimants has increased over the last several years and the rate of paid preparation has declined.

[70] The IRS sponsors programs that offer free tax return preparation services and counseling to seniors, individuals with low to moderate incomes, those with disabilities, and those for whom English is a second language; the programs are the Volunteer Income Tax Assistance (VITA) program and the Tax Counseling for the Elderly (TCE) program, both of which are staffed by specially trained volunteers.

Table 8. Number of Returns By Preparer Type and EITC Claim Status
Weighted Population Estimates, Annual Average, NRP TY 2006-2007[1]

Preparer Type[2]	Returns Not Claiming EITC			Returns Claiming EITC		
	Number of returns (millions)	Percent of all returns not claiming EITC (%)	Percent of those using a preparer where type is known (%)	Number of returns (millions)	Percent of all returns claiming EITC (%)	Percent of those using a preparer where type is known (%)
Self-prepared[3]	49.8	43%	--	6.9	29%	--
	0.2	0%		0.2	1%	
IRS/VITA/TCE[4]	2.4	2%	--	0.6	3%	--
	0.1	0%		0.1	0%	
Paid preparer	63.7	55%	--	16.2	68%	--
	0.2	0%		0.2	1%	
Attorney	0.7	1%	2%	0.0[6]	0%	0%
	0.1	0%	0%	0.0	0%	0%
CPA	19.0	16%	44%	1.5	6%	10%
	0.3	0%	1%	0.1	0%	1%
Enrolled agent	4.6	4%	11%	1.3	6%	9%
	0.2	0%	0%	0.1	0%	1%
Employee of taxpayer	0.0[6]	0%	0%	0.0[6]	0%	0%
	0.0	0%	0%	0.0	0%	0%
Friend or relative-paid	0.7	1%	2%	0.3	1%	2%
	0.1	0%	0%	0.0	0%	0%
National tax return preparation firm	6.0	5%	14%	5.0	21%	35%
	0.3	0%	1%	0.1	1%	1%
Unenrolled return preparer	12.1	10%	28%	6.3	26%	43%
	0.3	0%	1%	0.1	1%	1%
Preparer used, type unknown[5]	20.5	18%	--	1.8	8%	--
	0.4	0%		0.1	0%	
Total	115.9	100%	100%	23.7	100%	100%

Note: Figures may not sum due to rounding. Standard errors are presented below estimates.

1 Due to an inconsistency between TY 2006-2007 and TY 2008 in NRP data collection methods regarding preparer types, this analysis is limited to TY 2006 and TY 2007 only. Cases are reweighted to reflect the three-year population totals for EITC claimants and EITC non-claimants.

2 NRP auditors were asked to indicate the type of preparer for each return by selecting all that applied from the list of options shown in this table. For over 95 percent of the sample (unweighted), only one type of preparer was selected. For the remaining multiple selections, returns were grouped as follows. "Attorney" includes returns where CPA, enrolled agent, or unenrolled return preparer were also selected. "CPA" includes returns where enrolled agent or unenrolled return preparer were also selected. "Enrolled agent" does not include any multiple selections. "Employee of taxpayer" includes returns where attorney, CPA, national tax return preparation firm, or unenrolled return preparer was also selected. "Friend or relative-paid" includes returns where attorney, CPA, enrolled agent, or unenrolled return preparer was also selected, if the taxpayer reported compensating the preparer (cases where friends or family members were uncompensated are considered self-prepared for purposes of this table; see note 3, below). "National tax return preparation firm" includes returns where CPA, enrolled agent, or unenrolled return preparer was also selected. Finally, "unenrolled return preparer" does not include multiple selections.

3 Self-prepared returns include those where the taxpayer reported receiving uncompensated assistance from another individual. For the self-preparers claiming the EITC, 28 percent received this kind of informal assistance; for the self-preparers not claiming EITC, the percent receiving informal assistance is just 9 percent.

4 Returns in this category were mainly prepared at IRS-sponsored Volunteer Income Tax Assistance (VITA) or Tax Counseling for the Elderly (TCE) sites, although 3 percent were prepared or reviewed by IRS employees through other venues.

5 The majority of returns in this category are cases where the return was accepted as filed, so no detail on type of preparer was able to be collected during an audit. There is also a large number of audit non-participants in this group.

6 Less than 50 thousand.

This does not necessarily imply that taxpayers or other kinds of preparers are either less capable or more unscrupulous. There is good reason to think that substantial selection bias arises from taxpayers' choice of preparer. Taxpayers who are eligible for and seek services from IRS-sponsored programs may have simpler financial and family circumstances, making the correct determination of EITC easier. As a group, they may be more likely to be compliant than taxpayers who seek other forms of assistance. Thus, even with no difference in ability or desire to be compliant between the IRS-trained volunteers and other preparers, one might see this pattern. It is perhaps surprising that the frequency of overclaims and the dollar overclaim percentage are not closer to zero for the volunteer-prepared returns. One cannot discern from these data whether this is because the complexity of the credit makes it difficult to claim correctly even for volunteers, or whether taxpayers misrepresent their financial and family situations – even when seeking assistance from IRS-sponsored programs – in order to claim a (higher) credit.

Unenrolled return preparers are at the other end of the spectrum, with the highest frequency of error and overclaim percentages.[71] For these preparers, the likelihood of making an overclaim is 54 percent (higher estimate) or 49 percent (lower estimate) and the dollar overclaim percentage is 40 percent (higher estimate) or 33 percent (lower estimate). Again, due to the problem of selection bias, one cannot conclude anything about the relative ability or integrity of unenrolled return preparers without further research.

Nonetheless, the high frequency with which EITC claimants choose unenrolled return preparers, combined with their higher error rates, does suggest that substantially improving the quality of returns prepared by unenrolled return preparers would have an effect on overall measures of EITC compliance. The performance of the other preparer types should help shape expectations about how much improvement might be expected as a result of regulatory efforts, such as the proposal in the Administration's Fiscal Year 2015 Budget to explicitly authorize the IRS to regulate all paid tax return preparers.

[71] The group of paid preparers where preparer type is not known has even higher error rates, but this is accounted for by the high rate of audit non-participation in that group and other similar reasons.

Table 9. Overclaims and Underclaims on EITC Returns by Preparer Type
Weighted Population Estimates, Annual Average, NRP TY 2006-2007[1]
(Constant 2008 dollars)

Preparer Type[2]	Number of returns (millions)			Percent with over claim	Dollars (billions)			Dollar overclaim percentage by preparer type
	Under claims	Correct claims	Over claims		Under claims	Over claims	Total claims	
	Higher estimates							
Self-prepared[3]	0.4	3.3	3.3	47%	$0.1	$4.6	$12.0	39%
	0.0	0.1	0.1	1%	$0.0	$0.3	$0.5	2%
IRS/VITA/TCE[4]	0.1	0.4	0.2	26%	$0.0[7]	$0.1	$0.8	13%
	0.0	0.0	0.0	4%	$0.0	$0.0	$0.1	3%
Paid preparer	0.9	6.9	8.3	51%	$0.4	$14.1	$36.4	39%
	0.1	0.2	0.2	1%	$0.1	$0.5	$0.5	1%
Attorney	0.0[6]	0.0[6]	0.0[6]	35%	$0.0[7]	$0.0[7]	$0.1	28%
	0.0	0.0	0.0	17%	$0.0	$0.0	$0.0	17%
CPA	0.1	0.7	0.7	49%	$0.1	$0.8	$2.6	31%
	0.0	0.1	0.1	3%	$0.0	$0.1	$0.2	3%
Enrolled agent	0.1	0.6	0.6	46%	$0.0[7]	$0.8	$2.8	29%
	0.0	0.1	0.1	4%	$0.0	$0.1	$0.2	4%
Employee of taxpayer	0.0[6]	0.0[6]	0.0[6]	58%	$0.0[7]	$0.0[7]	$0.0[7]	5%
	0.0	0.0	0.0	40%	$0.0	$0.0	$0.0	3%
Friend/relative-paid	0.0[6]	0.1	0.1	37%	$0.0[7]	$0.1	$0.5	19%
	0.0	0.0	0.0	7%	$0.0	$0.0	$0.1	7%
National tax return preparation firm	0.3	2.5	2.2	44%	$0.2	$3.6	$11.8	30%
	0.0	0.1	0.1	2%	$0.0	$0.2	$0.4	2%
Unenrolled return preparer	0.4	2.5	3.4	54%	$0.2	$5.8	$14.5	40%
	0.0	0.1	0.1	1%	$0.0	$0.3	$0.4	2%
Preparer used, type unknown[5]	0.0[6]	0.5	1.3	72%	$0.0[7]	$3.0	$4.1	73%
	0.0	0.0	0.1	3%	$0.0	$0.2	$0.3	3%
Total - higher estimates	**1.3**	**10.7**	**11.7**	**49%**	**$0.5**	**$18.8**	**$49.1**	**38%**
	0.1	0.2	0.2	1%	$0.1	$0.5	$0.3	1%

Continued on next page

Table 9, Continued. Overclaims and Underclaims on EITC Returns by Preparer Type
Weighted Population Estimates, Annual Average, NRP TY 2006-2007[1]
(Constant 2008 dollars)

Preparer Type[2]	Number of returns (millions)			Percent with over claim	Dollars (billions)			Dollar overclaim percentage by preparer type
	Under claims	Correct claims	Over claims		Under claims	Over claims	Total claims	
	Lower estimates							
Taxpayer self-prepared[3]	0.4	3.8	2.7	39%	$0.1	$3.4	$12.0	28%
	0.0	0.1	0.1	1%	$0.0	$0.3	$0.5	2%
IRS/VITA/TCE[4]	0.1	0.5	0.1	20%	$0.0[7]	$0.1	$0.8	11%
	0.0	0.1	0.0	4%	$0.0	$0.0	$0.1	3%
Paid preparer	1.0	8.0	7.1	44%	$0.5	$10.5	$36.4	29%
	0.1	0.2	0.2	1%	$0.1	$0.4	$0.5	1%
Attorney	0.0[6]	0.0[6]	0.0[6]	35%	$0.0[7]	$0.0[7]	$0.1	28%
	0.0	0.0	0.0	17%	$0.0	$0.0	$0.0	17%
CPA	0.1	0.7	0.7	47%	$0.1	$0.7	$2.6	27%
	0.0	0.1	0.1	3%	$0.0	$0.1	$0.2	3%
Enrolled agent	0.1	0.7	0.6	42%	$0.0[7]	$0.7	$2.8	24%
	0.0	0.1	0.1	4%	$0.0	$0.1	$0.2	3%
Employee of taxpayer	0.0[6]	0.0[6]	0.0[6]	58%	$0.0[7]	$0.0[7]	$0.0[7]	5%
	0.0	0.0	0.0	40%	$0.0	$0.0	$0.0	3%
Friend/relative-paid	0.0[6]	0.1	0.1	37%	$0.0[7]	$0.1	$0.5	19%
	0.0	0.0	0.0	7%	$0.0	$0.0	$0.1	7%
National tax return preparation firm	0.4	2.9	1.8	36%	$0.2	$2.4	$11.8	20%
	0.0	0.1	0.1	2%	$0.0	$0.2	$0.4	2%
Unenrolled return preparer	0.4	2.8	3.1	49%	$0.2	$4.7	$14.5	33%
	0.1	0.1	0.1	2%	$0.0	$0.3	$0.4	2%
Preparer used, type unknown[5]	0.1	0.8	0.9	51%	$0.0[7]	$1.9	$4.1	47%
	0.0	0.1	0.1	5%	$0.0	$0.3	$0.3	6%
Total - lower estimates	**1.5**	**12.2**	**10.0**	**42%**	**$0.6**	**$14.0**	**$49.1**	**28%**
	0.1	0.2	0.2	1%	$0.1	$0.5	$0.3	1%

Note: Figures may not sum due to rounding. Standard errors are presented below estimates.

1 Due to an inconsistency between TY 2006-2007 and TY 2008 in NRP data collection methods regarding preparer types, the analysis in this table is for combined TY 2006 and TY 2007 only. Cases are reweighted to replicate the three-year population totals for EITC claimants and EITC non-claimants. As a result, the summary information presented in this table about overclaims, underclaims, and dollar overclaim percentages differs slightly from that describing the full three-year sample presented in Table 1.

2-5 See the notes under Table 8 for more description of these categories.

6 Less than 50 thousand.

7 Less than $50 million.

Conclusion

The NRP TY 2006-2008 data indicate that many aspects of EITC compliance are qualitatively unchanged from the 1999 Compliance Study, despite developments in the interim that include overall growth in the EITC program, new EITC-related legislation, and enhanced enforcement efforts by the IRS. The lower and higher estimates of the dollar overclaim percentage in TY 2006-2008 fall below and above the respective estimates from the 1999 Compliance Study, so that no change in overall compliance can be detected without making strong assumptions about the compliance behavior of audit non-participants. Income misreporting and qualifying child errors are the errors most frequently made and account for the highest dollar amounts of overclaims. The residency test is the most frequent qualifying child error. One notable change since the 1999 Compliance Study is that tiebreaker rules are no longer a major source of overclaims, due to provisions of EGTRRA.

This study provides some new information about return preparers and EITC errors. Of the EITC claimants who use a paid preparer where preparer type is known, 43 percent seek preparation services from unenrolled tax return preparers, who, as a group, have the highest error rates and overclaim percentages among known preparer types. Returns prepared by the IRS-sponsored programs VITA and TCE have the lowest error rates and overclaim percentages, but these constitute a very small percent of returns with EITC.

The data underlying the analysis in this report are from the IRS' National Research Program TY 2006-2008 1040 Studies. These data provide the opportunity to study many aspects of EITC compliance beyond the overview presented here. Research questions regarding errors that occur infrequently may require a larger sample in order to draw statistically valid conclusions, but with ongoing annual NRP 1040 Studies that continue to have an EITC subsample, this should be possible in the future.

References

Chetty, Raj, John Friedman, Peter Ganong, Kara Leibel, Alan Plumley, and Emmanuel Saez, 2011. "Taxpayer Response to the EITC: Evidence from IRS National Research Program," Presented at the Annual Meeting of the National Tax Association, November 18, 2011.

Holt, Stephen D. "The Earned Income Tax Credit at Age 30: What We Know," Brookings Institution, 2006.

Holtzblatt, Janet, and Janet McCubbin, "Whose Child Is It Anyway? Simplifying the Definition of a Child," *National Tax Journal*, Volume 56, September, 701-718, 2003.

Hotz, V. Joseph, and John Karl Scholz, "The Earned Income Tax Credit," In R. Moffitt,Ed., *Means-Tested Transfer Programs in the United States*. Chicago: The University of Chicago Press and NBER, 2003.

Internal Revenue Service, *IRS Data Book*, 2008-2011 and various years.

_____, National Research Program, *Weighting Documentation: IRS-1040 TY 2008*.

_____, *Statistics of Income--2008 Individual Income Tax Returns*.

_____, *Publication 596, Earned Income Credit*, 2006-2008 and various years.

_____, *Tax Year 2001 Federal Tax Gap*, February 2007, available at http://www.irs.gov/pub/irs-utl/tax_gap_update_070212.pdf

_____, *Report to Congress*: *IRS Tax Compliance Activities*, 2003.

_____, *Compliance Estimates for Earned Income Tax Credit Claimed on 1999 Returns*, February 28, 2002.

_____, *Study of EITC Filers for Tax Year 1994*, April 1997.

_____, *Statistics of Income--1975 Individual Income Tax Returns*.

Joint Committee on Taxation, *General Explanation of Tax Legislation Enacted in the 110th Congress*, March 2009.

Maag, Elaine, "Tax Simplification: Clarifying Work, Child, and Education Incentives," Urban Institute and Tax Policy Center, March 2011.

McCubbin, Janet, 2000, "EITC NonCompliance: The Determinants of the Misreporting of Children," *National Tax Journal*, Volume 53, No. 4, Part 2, December, 1135-1164.

National Taxpayer Advocate, *National Taxpayer Advocate 2011 Annual Report to Congress, Volume One,* 2011. Available at http://www.taxpayeradvocate.irs.gov/Media-Resources/FY-2011-Annual-Report-To-Congress-Full-Report

_____, *National Taxpayer Advocate 2004 Annual Report to Congress, Volume 2: Earned Income Tax Credit (EITC) Audit Reconsideration Study*, 2004. Available at http://www.irs.gov/pub/irs-utl/nta2004arcvol2interactive.pdf

Nellen, Annette, "Simplification of the EITC Through Structural Changes," in the *Study of the Overall State of the Federal Tax System and Recommendations for Simplification, Volume III: Academic Papers Submitted to the Joint Committee on Taxation*, April 2001.

Office of Management and Budget, *Circular A-123, Appendix C: Requirements for Effective Measurement and Remediation of Improper Payments*, April 14, 2011.

Plueger, Dean and Amy O'Hara, "Earned Income Tax Credit Participation Rate for Tax Year 2005." Presented at the 2009 IRS Research Conference and available at http://www.irs.gov/pub/irs-soi/09resconeitcpart.pdf

U.S. Census Bureau, *American Community Survey, Design and Methodology*, April 2009.

Wancheck, John and Robert Greenstein, "Earned Income Tax Credit Overpayment And Error Issues," Center on Budget and Policy Priorities, April 19, 2011.

Appendix

Comparison of 1999 and 2006-2008 EITC Compliance Studies

The two tables in this section of the Appendix provide a side-by-side comparison of some of the major findings from the 1999 Compliance Study and the 2006-2008 Compliance Study. Table A1 presents information about the frequency of particular errors, while Table A2 focuses on the dollars of overclaims accounted for by returns with certain errors or combinations of errors. In both tables, the figures for 1999 are taken from Table 2 of the 1999 Compliance Study. Figures for 2006-2008 are taken in part from Tables 3 and 4 of this report and include some additional work not otherwise presented. Because a major change since 1999 is the virtual elimination of tiebreaker errors (due to legislation), the tables include an additional breakdown of the 1999 errors that excludes the tiebreaker errors, for better comparison with 2006-2008. The rise of multiple errors reflected in the last row makes problematic any comparisons across time of individual error types.

Table A1. Frequency of errors in 1999 and 2006-2008

	1999			2006-2008	
	Number of returns (millions)	**Percent of returns**	**Percent of returns excluding tiebreaker errors**	**Number of returns (millions)**	**Percent of returns**
Total returns with error	9.3	100.0%		11.9	100.0%
Type of error unknown[1]	1.2	13.3%		3.6	30.3%
Type of error known	8.1	86.7%		8.4	70.6%
Subtotal: Type of error known	8.1	100.0%		8.4	100.0%
Qualifying child (QC) error only	1.3	16.3%	18.3%	1.3	15.4%
Income reporting errors only	3.4	41.7%	46.7%	4.3	51.0%
"Tiebreaker error" only	0.9	10.8%	--	0.0	0.5%
Filing status error only	0.7	8.7%	9.7%	0.4	5.0%
Filing status & QC error	0.3	3.4%	3.8%	0.1	1.7%
Errors corrected in processing only	0.7	9.1%	10.2%	0.2	2.9%
All other errors and combinations	0.8	10.1%	11.3%	2.0	23.6%

1 Taxpayer unwilling or unable to appear for audit

Table A2. Dollars of overclaims in 1999 and 2006-2008

	1999			2006-2008	
	Dollars (billions)	Percent	Percent of dollars excluding tiebreaker errors	Dollars (billions)	Percent
Total returns with error	$11.1	100.0%		$19.3	100.0%
Type of error unknown[1]	$2.1	18.5%		$7.9	40.8%
Type of error known	$9.1	81.5%		$11.4	59.2%
Subtotal: Type of error known	$9.1	100.0%		$11.4	100.0%
Qualifying child (QC) error only	$2.3	24.9%	30.0%	$3.0	26.3%
Income reporting errors only	$1.9	21.4%	25.9%	$2.9	25.3%
"Tiebreaker error" only	$1.6	17.2%	--	$0.1	0.7%
Filing status error only	$1.0	10.7%	12.9%	$0.8	7.4%
Filing status & QC error	$0.6	6.7%	8.1%	$0.3	2.9%
Errors corrected in processing only	$0.6	6.5%	7.8%	$0.1	1.3%
All other errors and combinations	$1.1	12.6%	15.3%	$4.1	36.2%

1 Taxpayer unwilling or unable to appear for audit

EITC-Related Errors and Contributions to Total EITC Overclaims

Table A3 provides summary estimates of how each type of error contributes to total overclaims. Because there is no simple way to disaggregate overclaims into separate error types due to returns with multiple errors, ranges of estimates are presented. These ranges incorporate two different approaches for handling multiple errors, as reflected by the estimates in Table 5a. Thus, the pairs of numbers do not reflect different assumptions about audit non-participants; rather, the two numbers reflect approaches that attribute more or fewer overclaim dollars to each error type based on the order of attribution.

Table A3. EITC-Related Errors as Percentage of Total Overclaim Dollars Weighted Population Estimates, Annual Average, NRP TY 2006-2008

Error type	Percentage of Total Overclaim Dollars		
Qualifying child error	42%	–	54%
Income misreporting (all types combined)	24%	–	32%
Self-employment income alone	*15%*	–	*23%*
AGI and investment income alone	*5%*	–	*8%*
Wage income alone	*3%*	–	*6%*
Filing status error	9%	–	17%
Error corrected in processing	3%	–	3%
Rules for all taxpayers claiming EITC	1%	–	5%
Tiebreaker error	1%	–	2%
Rules for taxpayers claiming EITC without children	0%	–	1%

Note: See Table 5a for more detail on error categories.