

An IRS-TPC Research Conference: Improving Tax Administration Through Research-Driven Efficiencies

*Papers Given at the
2015 IRS-Tax Policy Center Research Conference*

**Held at the Urban Institute
Washington, DC
June 18, 2015**

Compiled and edited by Alan Plumley*
Research, Analysis, and Statistics, Internal Revenue Service

Foreword

This edition of the *IRS Research Bulletin* (Publication 1500) features selected papers from the IRS-Tax Policy Center (TPC) Research Conference, “Improving Tax Administration Through Research-Driven Efficiencies,” held at the Urban Institute in Washington, DC, on June 18, 2015. Conference presenters and attendees included researchers from all areas of the IRS, officials from other government agencies, and academic and private sector experts on tax policy, tax administration, and tax compliance. In addition to those who attended in person, many participated live online, as the TPC broadcast video of the proceedings over the Internet. The videos are archived on their Website to enable additional participation. Online viewers participated in the discussions by submitting questions via e-mail as the sessions proceeded.

The conference began with welcoming remarks by Eric Toder, Co-Director of the Tax Policy Center and by Alain Dubois, the IRS Acting Director of Research, Analysis, and Statistics, who conveyed a welcome from IRS Commissioner John Koskinen. The remainder of the conference included sessions on innovative methods for improving resource allocation, taxpayer responses to rules and enforcement, improving tax administration by understanding taxpayer behavior, and helping taxpayers get it right. The keynote speaker was Dr. Lillian Mills, the Beverly H. and William P. O’Hara Chair in Business, Department of Accounting, The University of Texas at Austin. She offered some “Reflections on IRS/Academic Collaboration.”

We trust that this volume will enable IRS executives, managers, employees, stakeholders, and tax administrators elsewhere to stay abreast of the latest trends and research findings affecting tax administration. We anticipate that the research featured here will stimulate improved tax administration, additional helpful research, and even greater cooperation among tax administration researchers worldwide.

Acknowledgments

The IRS-TPC Research Conference was the result of preparation over a number of months by many people. The conference program was assembled by a committee representing research organizations throughout the IRS. Members of the program committee included: Alan Plumley, Ron Hodge, and Rahul Tikekar (National Headquarters Office of Research); Debi Kant (Office of Program Evaluation and Risk Analysis); Kevin Pierce (Statistics of Income); Ariel Wooten (Wage and Investment); Fran Cappelletti (Taxpayer Advocate); Charles Boynton (Large Business and International); Dawn Fitzelle (Tax Exempt and Government Entities), and Elaine Maag (Tax Policy Center). In addition, Blake Greene and Joanna Teitelbaum from the Tax Policy Center and Linda Addison from the IRS RAS Data Management Division oversaw numerous details to ensure that the conference ran smoothly.

This volume was prepared by Paul Bastuscheck, Lisa Smith, and Camille Swick (layout and graphics) and Beth Kilss (editor), all of the IRS Statistics of Income Division. The authors of the papers are responsible for their content, and views expressed in these papers do not necessarily represent the views of the Department of the Treasury or the Internal Revenue Service.

We appreciate the contributions of everyone who helped make this conference a success.

Janice M. Hedemann
Director, National Headquarters Office of Research
Chair, 2015 IRS-TPC Research Conference

An IRS-TPC Research Conference: Improving Tax Administration Through Research-Driven Efficiencies

Contents

Foreword.....	iii
1. Innovative Methods for Improving Resource Allocation	
❖ Estimating Marginal Revenue/Cost Curves for Correspondence Audits <i>Ronald H. Hodge II, Alan H. Plumley, Kyle Richison, Getaneh Yismaw, Nicole Misek, Matt Olson, and H. Sanith Wijesinghe</i>	3
❖ IRS Collectability Curve <i>Tom Beers, Carol Hatch, Joe Saldana, and Jeff Wilson</i>	17
❖ Analysis of Flow-Through Entities Using Social Network Analysis Techniques <i>Ashish Agarwal, Shannon Chen, Ririko Horvath, Larry May, and Rahul Tikekar</i>	45
2. Taxpayer Responses to Rules and Enforcement	
❖ Taxpayer Behavior Under Audit Certainty <i>Benjamin Ayers, Jeri Seidman, and Erin Towery</i>	57
❖ 2011–2012 Schedule M-3 Profiles of Schedule UTP Filers by IRC Section Cited <i>Charles Boynton, Portia DeFilippes, Ellen Legel, and Lisa Rupert</i>	61
❖ Individual Nonfilers and IRS-Generated Tax Assessments: Revenue and Compliance Impacts of IRS Substitute Assessments When Taxpayers Don't File <i>Saurabh Datta, Stacy Orlett, and Alex Turk</i>	113
3. Improving Tax Administration by Understanding Taxpayer Behavior	
❖ Behavioural Interventions in Tax Compliance: Evidence from Guatemala <i>Stewart Kettle, Marco Hernandez, Simon Ruda, and Michael Sanders</i>	141

❖ 2014 Taxpayer Choice Model (TCM): Designing Digital Communication Products To Reduce Phone and Mail Inventory <i>Courtney Rasey and Mackenzie Wiley</i>	167
❖ Examining and Addressing Taxpayer Expectations for Affordable Care Act (ACA) Automated Information <i>Ariel S. Wooten and Marisa E. McDaniels</i>	187
4. Helping Taxpayers Get It Right	
❖ Nontaxable Combat Pay Election and the Earned Income Tax Credit <i>Suzanne Gleason and Patricia K. Tong</i>	207
❖ Incorporating Supplemental Nutrition Assistance Program Data in Earned Income Tax Credit Administration: A Florida Case Study <i>Elaine Maag, Sara Edelstein, Devlin Hanson, Sarah Minton, Michael Pergamit, and Caroline Ratcliffe</i>	217
❖ IRS Preparer-Level Treatment Tests <i>Karen Masken</i>	231
5. Appendix	
❖ Conference Program	249



Innovative Methods for Improving Resource Allocation

**Hodge ♦ Plumley ♦ Richison ♦ Yismaw
Misek ♦ Olson ♦ Wijesinghe**

Beers ♦ Hatch ♦ Saldana ♦ Wilson

Agarwal ♦ Chen ♦ Horvath ♦ May ♦ Tikekar

Estimating Marginal Revenue/Cost Curves for Correspondence Audits

Ronald H. Hodge II, Alan H. Plumley, Kyle Richison, and Getaneh Yismaw,
Research, Analysis, and Statistics, Internal Revenue Service, *and*
Nicole Misek, Matt Olson, and H. Sanith Wijesinghe, MITRE Corporation

Tax agencies have long desired to allocate their resources so as to maximize the revenue they collect net of administrative costs.¹ It has been clear that net revenue is maximized when the marginal revenue/cost ratio is equalized across all potential activities; otherwise, net revenue could be increased by shifting resources from activities having low marginal revenue/cost to those that exhibit higher marginal revenue/cost. However, marginal revenue and marginal cost are usually not observed; they must be estimated, and that is often a challenge. As a result, many tax administrators rely on readily observable average revenue/cost metrics, which lead to sub-optimal outcomes.²

This paper provides empirical estimates of marginal revenue/cost functions for several important categories of correspondence audits of tax returns conducted by the Internal Revenue Service (IRS) for Tax Years 2006–2010, and explains the methodology we developed for making those estimates. We then use these curves to identify the allocation of resources among these audit categories for each of these historical years that would have maximized net direct revenue—suggesting that the same resources could have yielded an additional \$190 million of direct enforcement revenue if they had been allocated differently during these years.

1. Introduction

It is quite easy and intuitive to derive the average revenue per case, the average cost per case, and the average revenue-to-cost ratio (often called “Return on Investment,” or ROI) for a tax enforcement program. These metrics have the great advantage of being very straightforward to derive from observable data, and they have some use in managing enforcement programs and documenting results. However, none of these average measures provide the right basis for allocating scarce resources to the programs competing for those resources. That is, they cannot tell us *how much* of our resources we *should* devote to each program. In fact, devoting more of our budget to the programs exhibiting the highest average ROIs is often *not* the most cost-effective way to manage our resources. That is because our ultimate objective ought to be to maximize the net benefit we produce through our programs—that is, the total benefit minus the total cost—and the way to maximize net benefits is to equalize the *marginal* benefit/cost ratio across all programs; otherwise, net benefits could be increased by shifting resources from activities having low marginal benefit/cost to those that exhibit higher marginal benefit/cost.³ One of the benefits produced by a tax agency is the revenue it collects—whether paid voluntarily or in response to enforcement actions.⁴ This paper focuses solely on the revenue generated *directly*⁵ by one enforcement program—correspondence audits.⁶

¹ Sometimes in addition to other objectives.

² See the Government Accountability Office report, “TAX GAP: IRS Could Significantly Increase Revenues by Better Targeting Enforcement Resources,” GAO-13-151, December 2012.

³ More accurately, the uniform marginal benefit/cost ratio would apply to any program that is not otherwise constrained. For example, it would presumably be possible to divert resources away from any program whose marginal benefit/cost ratio is currently *less* than the optimal uniform ratio. But if its ratio is *greater* than the optimal uniform ratio, and there are short-term or permanent constraints preventing the program from being expanded to the point at which its ratio is lowered to the uniform ratio, then once that constraint becomes binding, the marginal benefit/cost ratio would effectively drop to zero. At that point, it would be counter-productive (at least while the constraint remains binding) to add resources to that program, though it would be helpful to make progress on relieving the constraint, if possible, cost-effectively.

⁴ We do not include any penalties or interest among the benefits since our objective is not to maximize them.

⁵ If we had estimates of the associated changes in voluntary compliance that are induced indirectly by that program throughout the entire population, those estimates could be added to the direct revenue estimates to represent the full benefit of the program. Although that is a long-term goal of our research, it is reasonable in the meantime to assume that these indirect effects are the same for each type of correspondence audit, so that they wouldn't influence the optimal allocation across correspondence audit categories.

⁶ Actually, we focus on just some of the major discretionary categories of correspondence audits.

The marginal revenue generated by a tax enforcement program generally declines as the level of effort (i.e., the budget) expended in that program increases.⁷ At a given budget, the marginal revenue is the change in revenue associated with changing that level of effort a little bit. Likewise, the marginal cost is the corresponding change in cost that produces that marginal revenue. Generally, marginal effects are not observed in isolation; they must be estimated. The research presented in this paper is (to our knowledge) the first to develop empirical estimates of marginal direct revenue/cost ratios for an IRS program, and it reflects the following key features:

- Analyzing the entire population of completed audits in given categories;
- Estimating for each audit the full administrative cost to conduct the audit, assess additional tax, and collect the tax due;
- Accounting for the amount of additional tax actually collected—not simply the amount proposed by the auditors or the amount formally assessed after appeal or litigation; and
- Ranking actual and potential audits in a given category according to their operational priority (risk).

The remainder of the paper proceeds as follows: Section 2 describes the correspondence audit program at the IRS; Section 3 outlines our estimation methodology; Section 4 presents our empirical results; Section 5 illustrates how these estimates could be used to improve the allocation of the budget within the correspondence audit program, surveying the volatility of the curves over time; and Section 6 concludes.

2. The Correspondence Audit Program

The IRS audits⁸ a small percentage of the tax returns filed in any given year. Most of these audits are initiated because the IRS perceives that the taxpayer may have misreported his tax obligation. There are three types of audits of individual tax returns, depending on the perceived nature of the noncompliance. The most complex issues require a face-to-face audit at the taxpayer's residence or place of business. Somewhat simpler cases are handled face-to-face in an IRS office. The simplest audits—typically requiring just documentation or responses to simple inquiries about a limited number of issues—are conducted through correspondence. Since Fiscal Year 2006, just over three-quarters of all of the individual income tax audits that were completed were conducted through correspondence, and they accounted for nearly 60 percent of the tax adjustments proposed by all individual income tax audits each year (Table 1).⁹ Correspondence audits tend to cost less—both in terms of IRS direct costs and also in terms of the burden placed on the taxpayers—so they are often quite cost-effective compared with face-to-face audits; hence the reliance on correspondence audits, which also underscores the importance of allocating our resources to the various categories of correspondence audits in the most cost-effective mix possible. Correspondence audits are categorized into “projects”, with each project devoted to one or a small number of specific lines on the tax return (e.g., specific sources of income, specific deduction items, a tax credit, or some combination of these), or to the special issues that are often present on returns filed late. Each project, then, consists of a fairly homogeneous set of returns. The returns within a given project that are actually selected for correspondence audit are even more homogeneous because they all must meet a set of project-specific selection criteria that suggest the possibility of misreporting. This first set of selection criteria (which are generally binary, in that they have a yes/no quality) typically produces more workload than the program has the resources to pursue. So, the program selects from among the returns that meet these initial screening criteria by prioritizing them according to a non-binary (i.e., continuously varying) indicator or computed risk score.

At the conclusion of any type of audit, the examiner proposes the final tax adjustment (if any) to the taxpayer. If the taxpayer agrees with this “recommended” adjustment, that amount is formally assessed and becomes legally due. If the taxpayer disagrees with some or all of the recommended amount, the taxpayer must challenge it either through administrative appeal with the IRS or through litigation, both of which will result in either an assessment of some or all of the recommended amount or an assessment of zero. Regardless of how or when an assessment is made, it is legally due. The assessed amount is often paid in full immediately (after a formal notice is sent to the taxpayer) or in installments, but sometimes collecting the tax due requires the IRS to devote additional resources to follow up with the taxpayer, and even to apply stronger collection tools such as liens or levies.

⁷ This is because most enforcement programs are successful in giving priority to cases that are at least somewhat more cost-effective to work than others. If the tax agency had no way of identifying in advance which cases would be more cost-effective than others, it would inevitably select cases randomly, whereupon the marginal revenue/cost ratio would be the same as the overall average—a constant.

⁸ Technically, the “audits” should be referred to as *examinations*, since they are not comprehensive audits of the taxpayer's tax obligation.

⁹ *Internal Revenue Service Data Book*, Table 9a; available each fiscal year on <http://www.irs.gov/uac/SOI-Tax-Stats-IRS-Data-Book>.

TABLE 1. Correspondence Examinations as a Share of all Examinations, FY 2006–2014

Fiscal Year	Returns Examined		Percent Correspondence	Recommended Additional Tax (\$ Millions)		Percent Correspondence
	Field	Correspondence		Field	Correspondence	
2006	302,785	981,165	76.4%	5,433	7,612	58.4%
2007	311,339	1,073,224	77.5%	6,357	9,348	59.5%
2008	310,429	1,081,152	77.7%	5,945	6,518	52.3%
2009	326,249	1,099,639	77.1%	7,145	7,796	52.2%
2010	342,762	1,238,632	78.3%	6,899	8,168	54.2%
2011	391,621	1,173,069	75.0%	5,947	8,705	59.4%
2012	359,750	1,122,216	75.7%	6,062	9,249	60.4%
2013	344,152	1,060,779	75.5%	5,594	8,455	60.2%
2014	291,643	950,836	76.5%	5,026	6,859	57.7%

NOTE: Roughly 500,000 correspondence audits relate to the Earned Income Tax Credit.
Source: IRS Data Books, Fiscal Years 2006–2014.

3. Estimation Methodology

The ultimate objective of any tax enforcement program ought to be to maximize its net benefits, which we can simplify in our context to mean maximizing net direct revenue (the total revenue actually collected minus the total cost of the audit, appeals, litigation, and collection processes applied to collect that revenue).¹⁰ It's not cost-effective to maximize net recommendations or net assessments; if the costs incurred on an audit produce a recommended or assessed tax change, but they don't actually produce net revenue collected, then the costs expended on that case were wasted.¹¹ Therefore, the methodology presented in this paper estimates marginal *revenue* (the amount of tax ultimately collected or paid) and marginal *ultimate* cost (the cost of the full audit, appeals, litigation, and collection life cycle of the cases).

Our methodology followed nine basic steps:

1. **Segment the Population:** We divide all completed audits for a given year into groups that are somewhat homogeneous with respect to taxpayer and agency behavior. For correspondence audits, the projects were the appropriate groups since each project reflects a unique compliance behavior and has its own criteria for selecting tax returns for audit. For some projects, we further segment the returns into High Income and Low Income sub-groups.
2. **Identify Revenue Collected:** For each audit, we identify the amount of tax¹² that was eventually collected. Since this ultimate disposition of the case can take some time, it is best to focus on a year for which the collection process has run its course. Fortunately, few correspondence audits take an extended amount of time to fully close.
3. **Estimate Cost:** For each audit, we also estimate the total cost of the entire examination, appeals, litigation, notice, and collection process that was incurred to collect the revenue for that case. IRS enforcement data include the number of hours (by type of employee) spent in the examination, appeals, and litigation steps (if applicable),

¹⁰ More generally, the benefits should include such things as the potential increase in revenue paid voluntarily in the general population due to the *indirect* (possibly deterrent) effect of the enforcement, and the costs should include any unnecessary monetary and non-monetary compliance costs borne by taxpayers and third parties in connection with the enforcement (routine, necessary post-filing compliance costs borne by taxpayers may serve as a deterrent, prompting taxpayers to avoid enforcement, which is a positive outcome that shouldn't be treated as a cost that nets against benefits). However, those non-direct benefits and costs are very difficult to quantify, and it is not clear how the various components of benefits and costs should be weighted relative to each other. Nonetheless, direct revenue and costs are the foundational components, and a tax enforcement program could maximize net direct revenue subject to reasonable constraints intended to account for the missing components from the ideal objective function.

¹¹ Some people might argue that conducting such an audit may well have an indirect deterrent value, promoting better voluntary compliance in the general population, but that would likely be more true if the case produced net revenue. And even if two cases produced the same indirect effect, one that produced no net direct revenue would clearly be less advantageous than one that did produce net direct revenue. Others may contend that not auditing returns we expect are understating their tax, but that we anticipate will not produce any net revenue, might prompt some taxpayers to display characteristics that could plausibly be expected to make them appear to be unlikely to pay a tax assessment without a lot of additional effort on the part of the IRS. It's possible that assessing less than the full recommended amount in the appeals or litigation process, or unsuccessful collection efforts might deter such taxpayers (or others in the population) from driving up the costs of IRS enforcement programs or reducing the collectability of enforcement assessments, but it seems more probable that the reverse is true as taxpayers may perceive that they can end up paying much less than the amount originally recommended by appealing the adjustment and dragging out the collection process.

¹² We excluded any interest and penalties that were paid, since we ought not let these become an incentive for us to delay the collection of tax. We also excluded non-enforcement revenue where it existed.

so it was a straightforward matter to estimate the cost of these steps using known hourly costs for each type of enforcement employee. Likewise, we have data indicating the number and type of automated notices sent to the taxpayer demanding payment of the final amount that was assessed, so it was easy to estimate that component of the cost. However, since the hours spent in the collection process are not captured separately for each tax return, we had to estimate that component of the overall cost for each audit. Fortunately, very few correspondence audits require such collection efforts. So for those that did, we applied the overall average cost per case of the collection process, accounting for the Automated Collection System and the Field Collection functions separately, depending on which of these a given audit case required.

4. Sort the Observations According to Priority: All of the returns in the database of completed audits met the basic project-specific selection criteria, but they differed with respect to a non-binary (i.e., continuously varying) indicator, as discussed in Section 2. Returns with a high value for the indicator were selected for audit, but the same indicator implicitly assigns a priority to each return, since those with the lowest value would not have been selected for audit had resources been tighter. So, we sorted the completed audits in declining order of the relevant indicator.
5. Compute Cumulative Revenue and Cumulative Cost: For each tax return in a given segment (i.e., project) in the given year, we computed the total revenue that would have been collected had that audit been the lowest priority audit to have been conducted. The revenue collected from each successive audit in the sorted file was added to the cumulative revenue up through the previous audit in the sorted file. The same procedure was used to compute the cumulative cost for each audit in the sorted file.
6. Plot Cumulative Revenue vs. Cumulative Cost: For each segment for a given year, we construct a plot of cumulative revenue vs. cumulative cost. This typically exhibits a slightly curved pattern, with the higher priority audits (on the left of the graph—at low levels of cumulative cost) having a higher revenue/cost than those with lower priority (on the right of the graph).
7. Fit a Curve: We used regression analysis to fit a curve through the observations in the plot constructed above. We found that a simple quadratic specification fit very well for most of the projects (which has simplifying benefits for the next step). In this case, we chose the following functional form:

$$\text{Cumulative Revenue} = a(\text{Cumulative Cost}) + b(\text{Cumulative Cost})^2 + \varepsilon \quad (1)$$

where a and b are the parameters to be estimated and ε is a disturbance term. No constant (intercept) term was included since cumulative revenue should be zero when cumulative cost is zero.

For some projects, a Power curve provided a better fit to the data. This has the following functional form:

$$\begin{aligned} \text{Cumulative Revenue} &= e^a (\text{Cumulative Cost})^b + \varepsilon_0 && \text{or} \\ \ln(\text{Cumulative Revenue}) &= a + b \ln(\text{Cumulative Cost}) + \varepsilon && (2) \end{aligned}$$

where a and b are again the parameters to be estimated and ε is a disturbance term.

8. Derive Marginal Revenue/Cost: We can easily derive Marginal Revenue/Cost (MR/MC) as a function of Cumulative Cost (i.e., budget) by taking the first derivative of equations (1) and (2) with respect to Cumulative Cost. The results are:

$$\text{Quadratic: MR/MC} = a + 2b(\text{Cumulative Cost}) \quad (3)$$

$$\text{Power: MR/MC} = e^a b(\text{Cumulative Cost})^{b-1} \quad (4)$$

Since we specified equation (1) as a quadratic, equation (3) is simply a straight line, with intercept a and a slope of $2b$. Hence, a should be positive, as it represents the revenue/cost of the first (highest priority) audit. And since we expect declining MR/MC, b should be negative. The value of equation (3) at a given level of Cumulative Cost is the slope of the fitted curve (equation (1)) at that level of resources. Given the nature of Equation (2), Equation (4) is nonlinear, approaching the MR/MC-axis asymptotically, and approaching a horizontal line (relatively constant MR/MC) at high levels of cost. That is, for declining MR/MC, a should be positive, and b should be between 0 and 1.

9. Derive the Optimal Allocation:¹³ As indicated earlier, the allocation that maximizes net direct revenue is the one in which the marginal revenue/cost ratio is the same across all segments competing against each other for resources. A straightforward way to do this is to transform equations (3) and (4) to express Cumulative Cost (i.e., budget) as a function of the marginal revenue/cost (MR/MC) ratio, as follows:

$$\text{Quadratic: Budget} = (\text{MR/MC} - a) / 2b \tag{5}$$

$$\text{Power: Budget} = \left[\frac{\text{MR/MC}}{be^a} \right]^{1/b-1} \tag{6}$$

We can then use equations (5) and (6) to calculate the budget required for each segment as a function of a common MR/MC ratio (using the estimated parameters *a* and *b* relevant to each segment), and therefore the total budget across all segments as a function of the common MR/MC ratio. The optimal allocation is the one in which the total budget estimated in this way is the same as the budget actually available (i.e., the overall cumulative cost).¹⁴

4. Empirical Results

We applied this methodology to the following seven correspondence audit project categories:

TABLE 2. Correspondence Audit Categories Studied

Project	Issue(s) Addressed
C1	A sole proprietor issue reported on Schedule C
A1-Lo	Various items claimed on Schedule A (Lo means low income taxpayers) (Hi means high income taxpayers)
A1-Hi	
A2	
A3	
A4	
O	Another Form 1040 issue

These projects account for many—but not all—of the correspondence audits conducted during these years as an illustration of the methodology. Table 3 provides an overview of the data for these seven projects for Tax Years 2006–2010, the most recent years available for analysis. Notice that the number of audits conducted and the revenue and cost varied significantly across these projects. Project O had the highest volumes and the highest average revenue/cost ratio, but it's not clear from these average metrics how to allocate the budget across these seven project categories. That will depend on the *marginal* revenue/cost functions, which we estimated using the methodology described in Section 3. Figures 1 through 7 show the basic plots of cumulative revenue vs. cumulative cost using the raw data, as well as the fitted curves through those data and the overall average revenue/cost lines for Tax Year 2006. Notice that they all display the expected curvature (diminishing marginal revenue/cost), although it is typically rather modest (presumably due to the rather homogeneous nature of the projects and the simple method used to assign priority to the cases).

¹³ We use the term “optimal” in the narrow context in which our simplified objective is to maximize net direct revenue (i.e., the revenue collected directly from the audits minus the full administrative cost to identify, assess, and collect that revenue). This may not be optimal in the ultimate context in which we account for all other benefits and costs to the IRS, taxpayers, and third parties, but this is a necessary starting point. Moreover, it is possible to impose constraints in this simple framework that limit the extent to which segments (i.e., projects) can be expanded or contracted. Some of these constraints may be known and quantifiable workload or resource limitations, while others may be subjective rules (such as minimum coverage constraints) that attempt to account for the likelihood and impact of indirect benefits.

¹⁴ This procedure would need to be modified to account for any constraints imposed on the expansion or contraction of the audit program in any of the segments. In the absence of such a constraint, a segment whose marginal revenue/cost ratio is below the optimal ratio at all budget levels would not have any audits conducted at all, and a segment whose marginal revenue/cost ratio is above the optimal ratio even for the last potential audit in the population would be given the budget to conduct audits on all possible returns.

TABLE 3. Summary Data for Selected Correspondence Audit Projects, Tax Years 2006–2010

Project	Tax Period	Number of Exams Completed	Total Revenue Collected (\$M)	Total Administrative Cost (\$M)	Average Revenue/Cost	Average Revenue/Exam	Average Cost/Exam
C1	2006	10,706	\$28.2	\$4.0	7.0	\$2,636	\$374
	2007	24,740	\$68.8	\$8.7	7.9	\$2,781	\$351
	2008	19,114	\$48.2	\$5.3	9.1	\$2,522	\$277
	2009	23,185	\$61.6	\$7.4	8.3	\$2,657	\$321
	2010	29,495	\$40.1	\$5.9	6.8	\$1,358	\$200
A1 Low Income	2006	67,621	\$159.4	\$13.4	11.9	\$2,357	\$198
	2007	58,615	\$126.6	\$12.0	10.6	\$2,161	\$204
	2008	27,803	\$64.4	\$5.2	12.3	\$2,317	\$188
	2009	41,528	\$72.0	\$8.1	8.9	\$1,734	\$196
	2010	31,614	\$36.9	\$5.7	6.5	\$1,169	\$181
A1 High Income	2006	9,895	\$19.3	\$2.4	8.1	\$1,947	\$240
	2007	9,459	\$27.1	\$2.7	10.0	\$2,870	\$286
	2008	7,813	\$23.9	\$2.5	9.5	\$3,063	\$323
	2009	3,428	\$2.1	\$0.7	3.2	\$623	\$196
	2010	3,080	\$3.9	\$0.6	6.0	\$1,261	\$210
A2	2006	30,541	\$67.7	\$7.1	9.5	\$2,216	\$233
	2007	27,639	\$62.2	\$6.8	9.2	\$2,252	\$245
	2008	2,480	\$5.1	\$0.4	14.2	\$2,070	\$146
	2009	8,838	\$15.1	\$2.0	7.6	\$1,714	\$226
	2010	2,135	\$3.3	\$0.5	6.6	\$1,540	\$233
A3	2006	21,298	\$22.3	\$4.3	5.2	\$1,048	\$201
	2007	10,961	\$9.2	\$2.0	4.7	\$840	\$179
	2008	21,011	\$16.2	\$3.4	4.7	\$771	\$163
	2009	13,043	\$8.3	\$1.6	5.3	\$634	\$119
	2010	12,768	\$7.3	\$1.8	4.1	\$570	\$139
A4	2006	13,919	\$37.5	\$3.0	12.6	\$2,696	\$214
	2007	41,069	\$86.8	\$9.1	9.6	\$2,114	\$220
	2008	23,885	\$63.1	\$5.6	11.2	\$2,641	\$235
	2009	17,942	\$48.9	\$4.9	10.1	\$2,723	\$271
	2010	31,799	\$50.7	\$5.9	8.6	\$1,595	\$184
O	2006	202,575	\$287.1	\$37.0	7.8	\$1,417	\$183
	2007	116,555	\$177.4	\$21.2	8.4	\$1,522	\$182
	2008	105,503	\$96.3	\$18.0	5.4	\$913	\$170
	2009	78,994	\$79.2	\$12.3	6.4	\$1,003	\$156
	2010	29,711	\$15.6	\$3.7	4.2	\$525	\$126

FIGURE 1. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code C1, Tax Year 2006

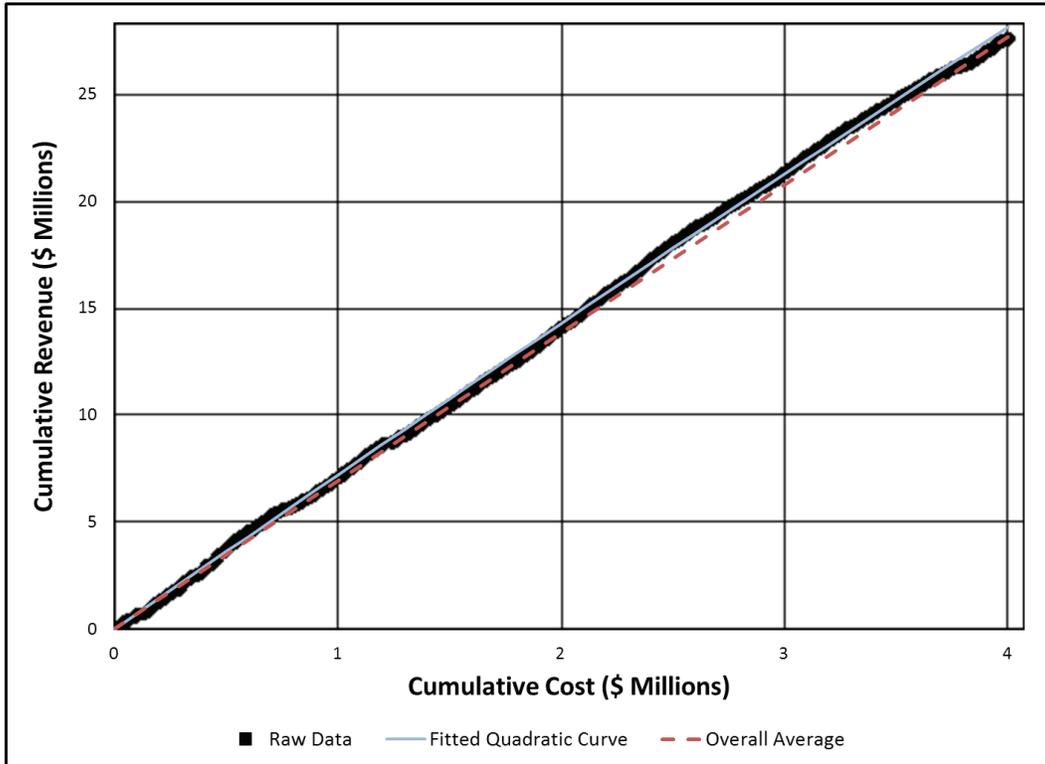


FIGURE 2. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code A1-Low Income, Tax Year 2006

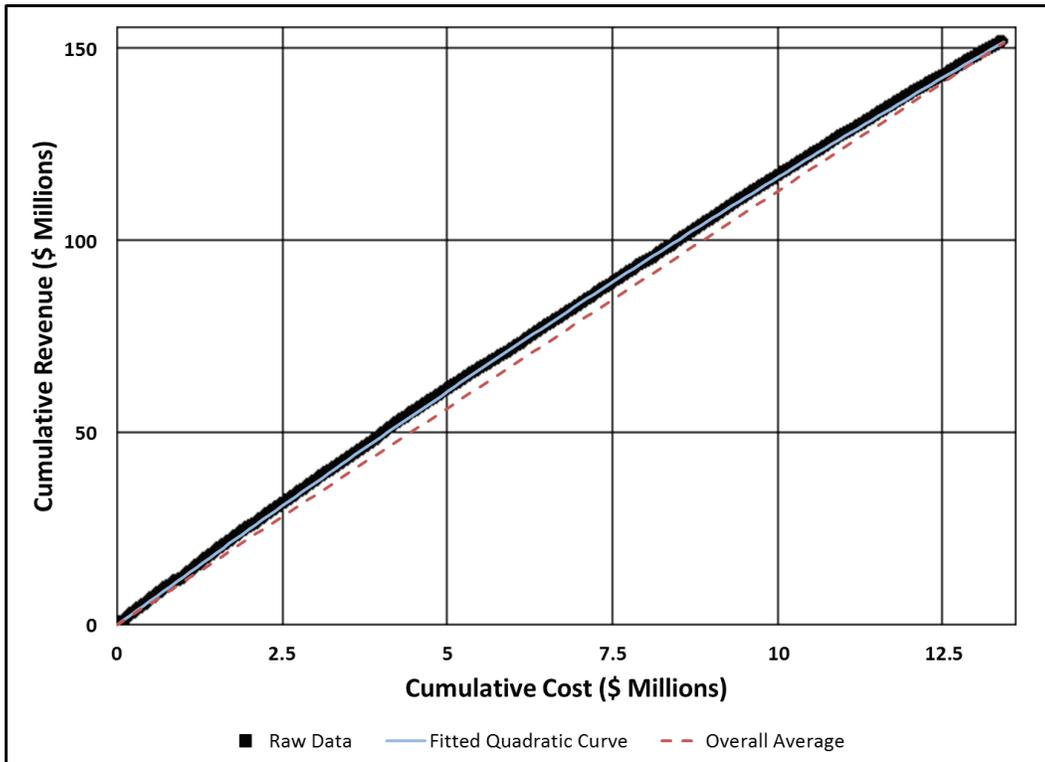


FIGURE 3. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code A1-High Income, Tax Year 2006

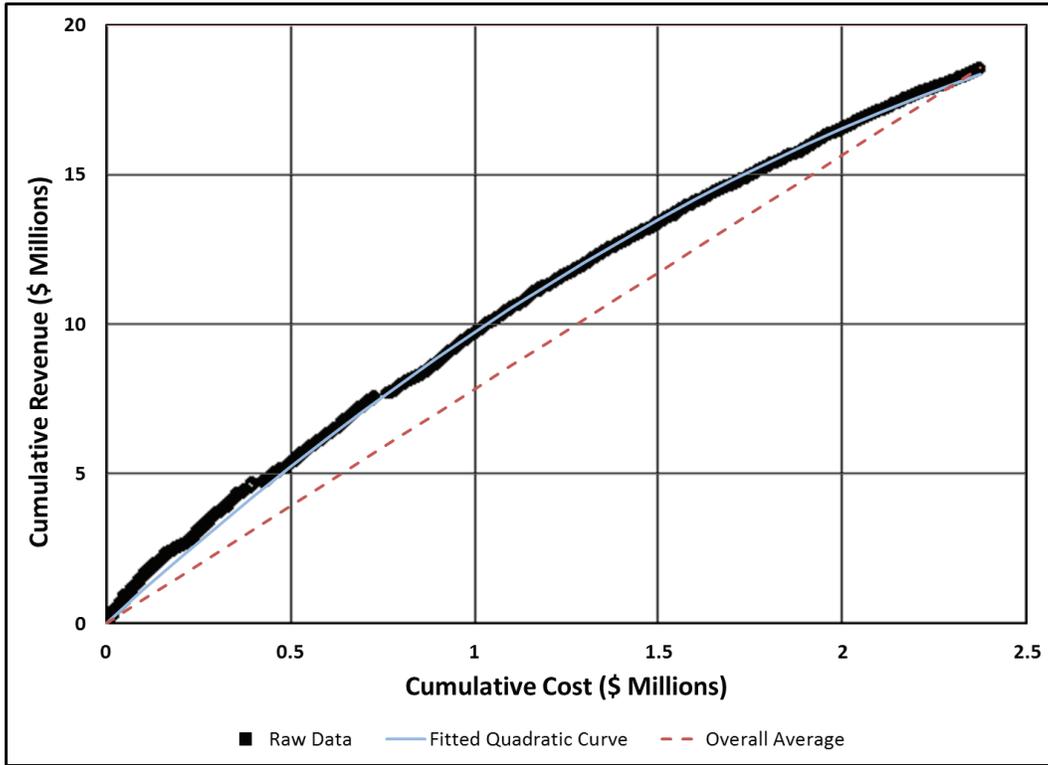


FIGURE 4. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code A2, Tax Year 2006

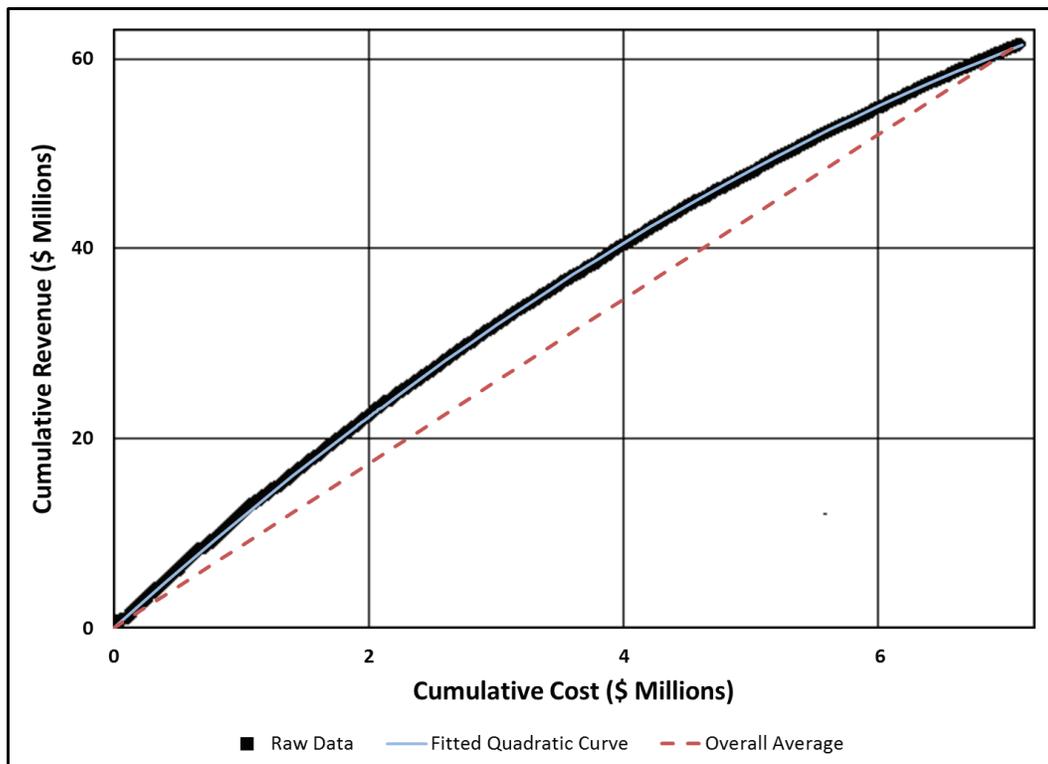


FIGURE 5. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code A3, Tax Year 2006

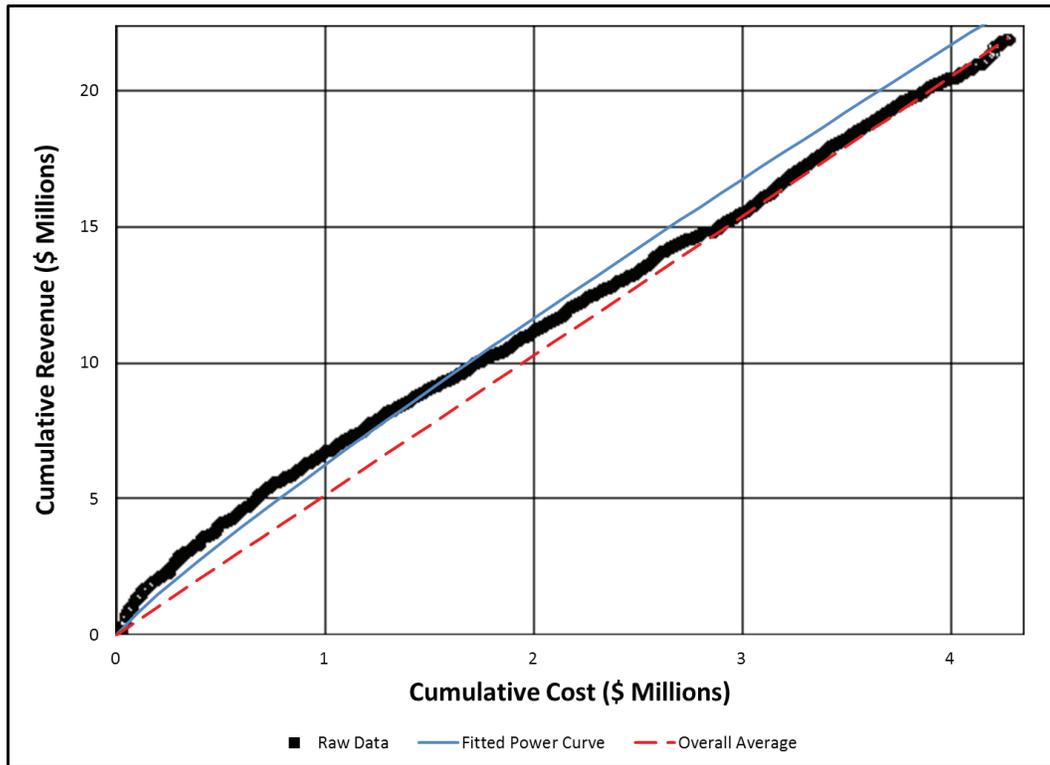


FIGURE 6. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code A4, Tax Year 2006

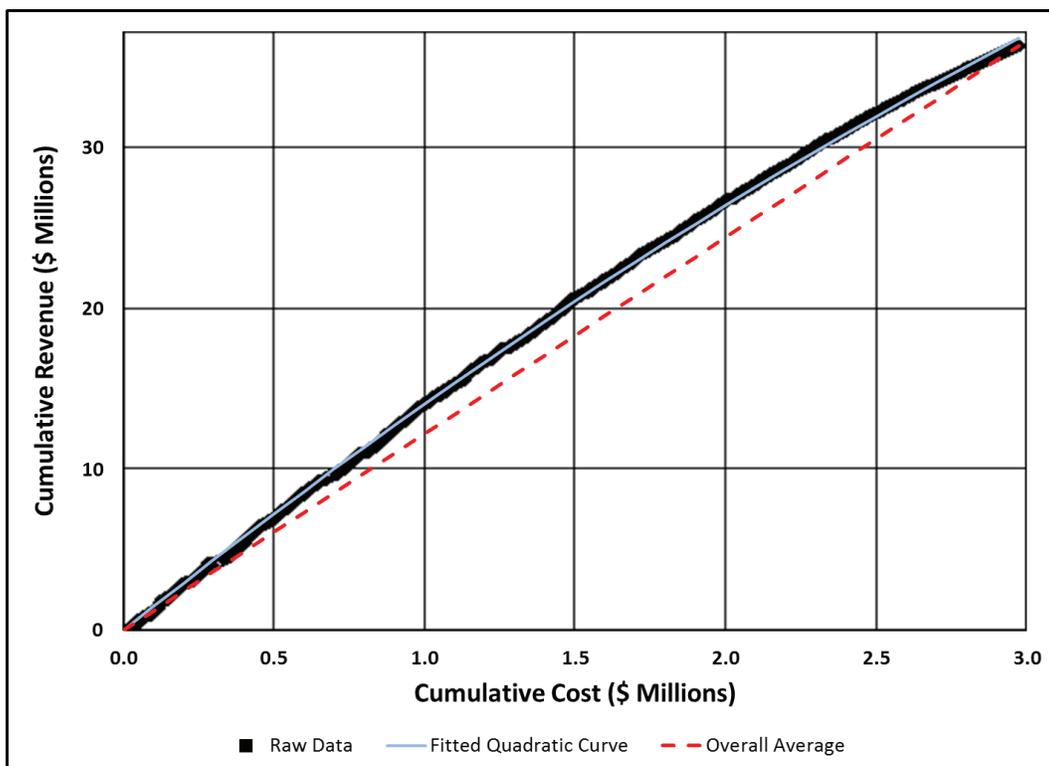
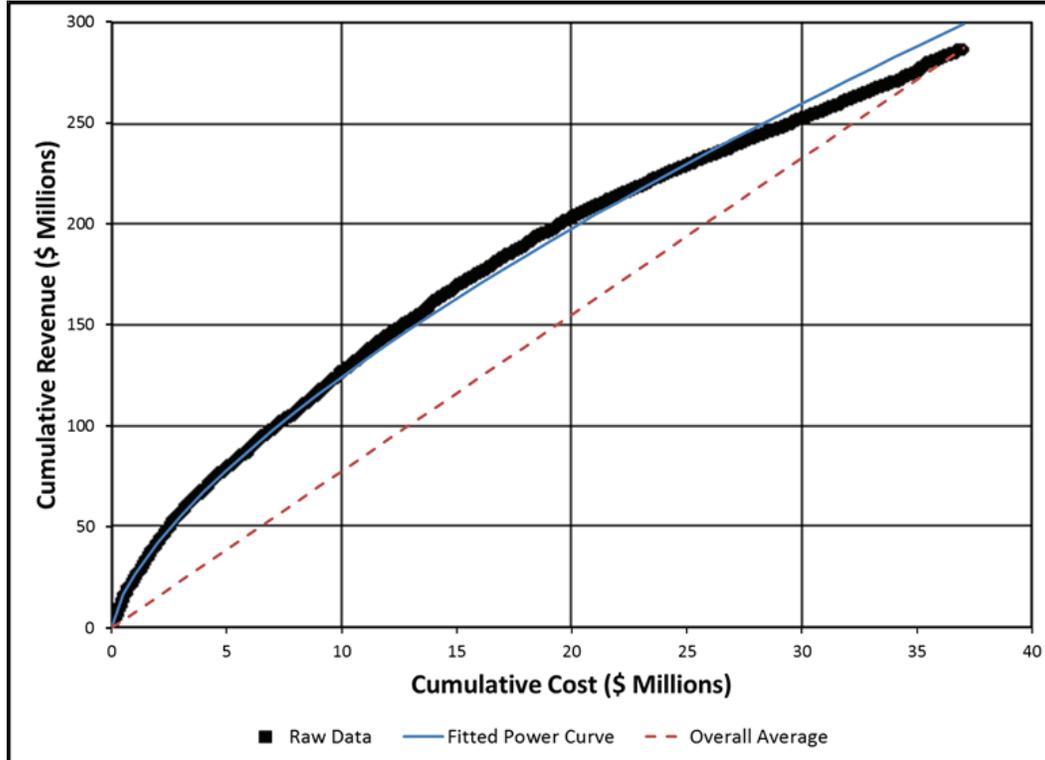


FIGURE 7. Plot of Sorted Data and Average and Fitted Curves of Revenue vs. Cost, Correspondence Audit Project Code O, Tax Year 2006



Notice that some of the curves (particularly Project Codes A2 and O) demonstrate strongly declining marginal revenue/cost (as reflected in their curvature), while others (e.g., Project Codes C1 and A1-Low) differ very little from the overall average line. In all of the projects, however, it may be possible to identify an improved prioritization metric, which would allow us to select much more cost-effective returns to audit than at present, resulting in more strongly declining revenue/cost and much higher overall revenue.¹⁵

The regression results for the fitted curves for Tax Year 2006 are presented in Table 4. All of the parameters are highly significant (due to the large number of observations), and have the expected signs. The results for the other years are similar, except that the parameter b for Project Code A1-Low was slightly positive in Tax Year 2008 (indicating a slightly *increasing* marginal revenue/cost due to the weakness of the prioritization variable).

5. Resource Allocation Implications

Applying equations (3) and (4), we can easily derive the marginal revenue/cost for each project as a function of the cumulative cost (or budget) allocated to it. This is merely the slope of the fitted curve at each point. Figure 8 plots all seven marginal curves together, and Table 5 provides the numerical results. Given a total budget of \$71.2 million, the maximum net direct revenue that could have been generated by these seven projects was \$597.4 million or \$63.5 million more than was actually generated. This outcome would have been produced if the budget had been allocated to these projects such that they all shared a marginal revenue/cost of 6.51. This allocation would have shifted resources from Projects O, A1-Hi, A2, and A3 to Projects C1, A1-Lo, and A4. For the most part, the employees who audit returns in these projects are interchangeable, and there appear to be ample returns that could have been audited in Project A1-Lo, so this allocation appears to have been feasible.¹⁶ It is not certain that had Project A1-Lo been expanded to this extent the estimated marginal revenue/cost function would have continued along the same straight line, but the strong relationship between revenue and cost suggests that this assumption is a strong basis for resource allocation.

¹⁵ This will be the subject of future research.

¹⁶ If this were not true in the short term, steps could be taken over time to move toward the optimal allocation, to the extent this is cost-effective.

TABLE 4. Regression Results for the Estimated Curves, Tax Year 2006

(Dependent Variable = Cumulative Revenue, CR)

Project	Functional Form (where C = cumulative cost)	a	b	Adjusted R-squared
C1	$CR = aC + bC^2$	7.23029 (2021.77)	-5.21147E-8 (-45.91)	0.9999
A1-Lo	$CR = aC + bC^2$	12.70311 (16684.8)	-1.01473E-7 (-1399.8)	1.0000
A1-Hi	$CR = aC + bC^2$	11.17570 (2228.35)	-1.45E-6 (-534.09)	0.9998
A2	$CR = aC + bC^2$	12.09733 (13223.5)	-4.87995E-7 (-2974.0)	1.0000
A3	$CR = e^a C^b$	3.23558 (117.31)	0.89817 (465.98)	0.9107
A4	$CR = aC + bC^2$	14.91579 (3639.92)	-8.55108E-7 (-495.50)	0.9999
O	$CR = e^a C^b$	7.83004 (2623.44)	0.67055 (3740.45)	0.9857

NOTE: t-statistics in parentheses

FIGURE 8. Estimated Marginal Revenue/Cost Curves, Selected Projects, Tax Year 2006

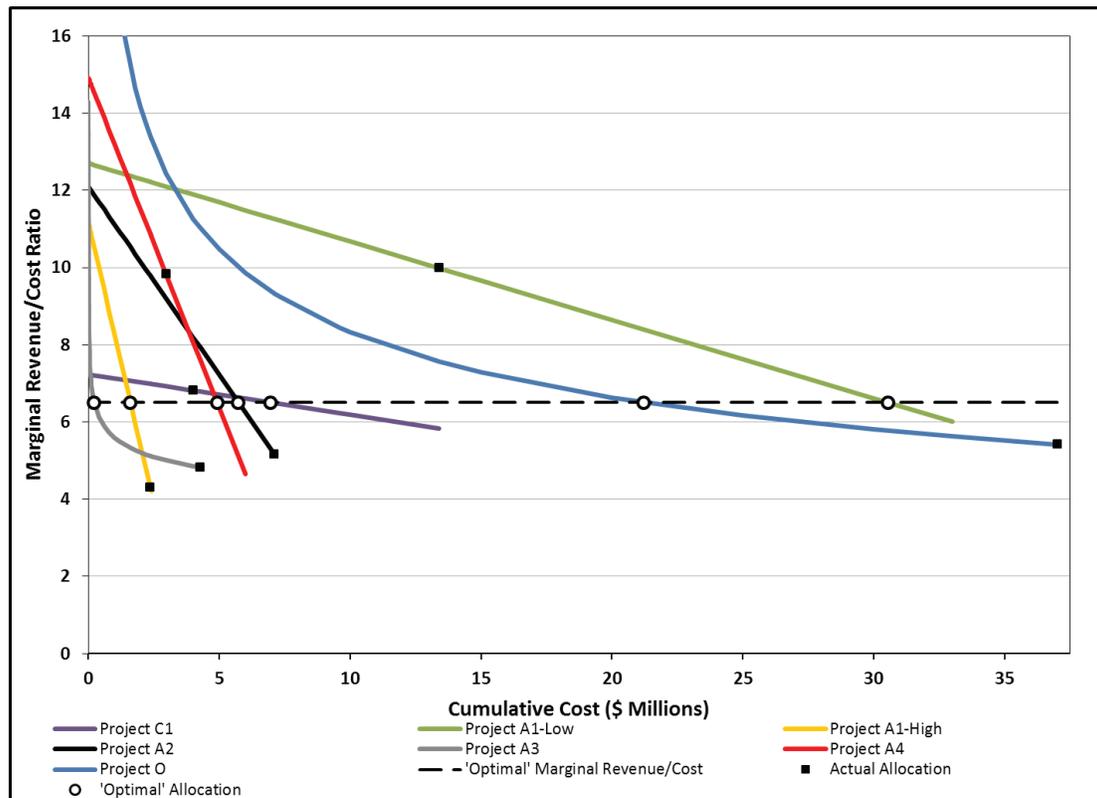


TABLE 5. Actual vs. Optimal* Revenue and Cost, Tax Year 2006 (\$ in Millions)

Project	Actual Allocation					Optimal* Allocation					Change in Cost	Change in Revenue
	Total Revenue	Total Cost	Net Revenue	Average Rev/Cost	Marginal Rev/Cost	Total Revenue	Total Cost	Net Revenue	Average Rev/Cost	Marginal Rev/Cost		
C1	\$27.6	\$4.0	\$23.6	6.90	6.81	\$47.7	\$7.0	\$40.8	\$6.87	6.51	\$2.9	\$20.1
A1-Lo	\$151.9	\$13.4	\$138.5	11.33	9.98	\$293.3	\$30.5	\$262.7	\$9.60	6.51	\$17.1	\$141.4
A1-Hi	\$18.6	\$2.4	\$16.2	7.83	4.29	\$14.2	\$1.6	\$12.6	\$8.84	6.51	-\$0.8	-\$4.3
A2	\$61.7	\$7.1	\$54.6	8.67	5.16	\$53.3	\$5.7	\$47.6	\$9.30	6.51	-\$1.4	-\$8.4
A3	\$21.9	\$4.3	\$17.6	5.12	4.82	\$1.6	\$0.2	\$1.4	\$7.24	6.51	-\$4.1	-\$20.3
A4	\$36.3	\$3.0	\$33.4	12.22	9.83	\$52.7	\$4.9	\$47.8	\$10.71	6.51	\$1.9	\$16.3
O	\$287.1	\$37.0	\$250.0	7.75	5.41	\$205.7	\$21.2	\$184.5	\$9.70	6.51	-\$15.8	-\$81.4
Total	\$605.1	\$71.2	\$533.9	8.50	N/A	\$668.6	\$71.2	\$597.4	\$9.39	6.51	\$0.0	\$63.5

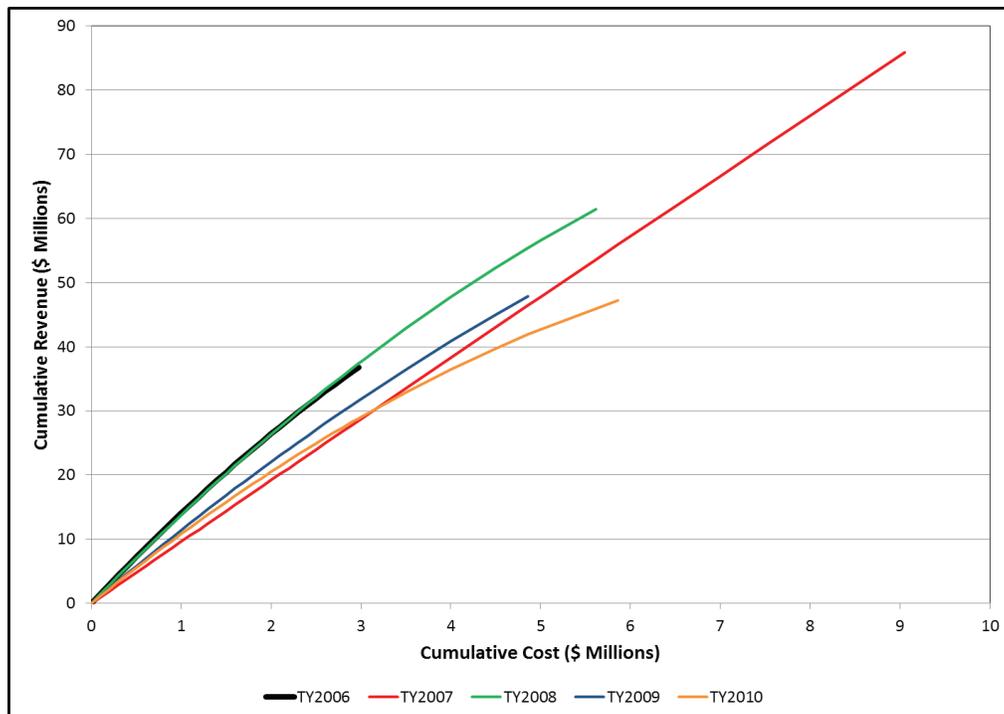
*Optimal only in the sense of maximizing direct revenue in an unconstrained setting, without accounting for indirect effects

Table 5 also shows that, at least for these projects, the average revenue/cost ratios under the actual allocation did not have the same ranking as the corresponding marginal revenue/cost ratios, and the optimal allocation was not at all proportional to or consistent with those average revenue/cost ratios. Most importantly, if the same resources had been allocated according to the *average* ratios, the largest share of the budget would presumably have been devoted to Project A4 (since its ratio of 12.22 was highest), but the optimal allocation (unlike the actual allocation) devotes the largest share of the budget to Project A1-Lo, which also had the highest marginal revenue/cost ratio in the actual allocation. The key is that without the marginal revenue/cost framework, even though decision-makers might want to shift resources from Projects C1 and A3 (given their low average ratios), they would not know *how much* of each budget to redirect to Projects A4 or A1-Lo; indeed they likely would have taken resources *from* Project C1 instead of *adding* resources to it, and they likely would have favored Project A4 over Project A1-Lo (given their relative average ratios).

Year-to-Year Fluctuations

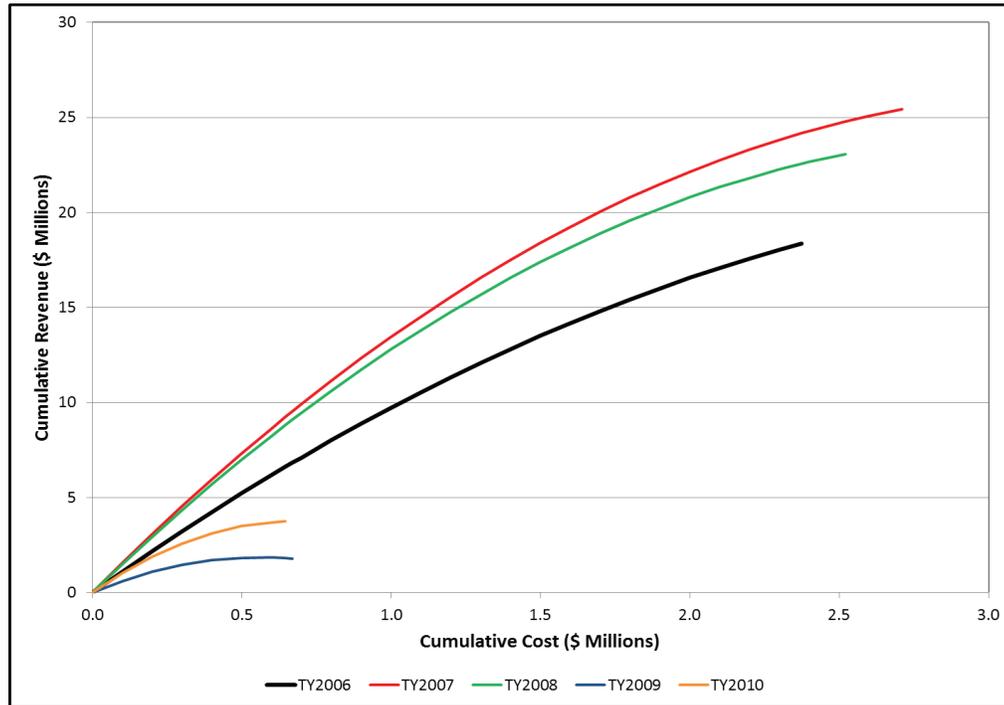
Ideally, the relationship between revenue and cost is relatively stable from year to year. However, this seems not to have been true in general. For example, Figure 9 shows how the cumulative revenue vs. cumulative cost plot for Project A4 varies across the five years of this study. While there was some consistency, there was also significant variation.

FIGURE 9. Cumulative Revenue vs. Cumulative Cost Curves, Project A4, Tax Years 2006–2010



There was even more variation in Project A1-Hi, as shown in Figure 10. In both cases, the budget (total cost) varied widely across these years, but the curves themselves varied, too.¹⁷ One might expect that the relationship between cumulative revenue and cumulative cost would remain fairly stable, and that a change in the budget would represent moving to a different location on the curve. But this appears not to be the case, making it challenging to use historical audit results to guide resource allocation in the current year.

FIGURE 10. Cumulative Revenue vs. Cumulative Cost Curves, Project A1-Hi, Tax Years 2006–2010



The variation over time is also seen in Table 6, which presents the change in revenue and the change in cost (budget allocation) for each project and year had the budget been allocated “optimally” with perfect knowledge of the marginal revenue/cost curves for the year in question. Obviously, the changing curves produce a very different “optimal” mix of projects each year. Further evidence of the time-sensitive nature of these relationships is given in Table 7, which corresponds to the estimates summarized in Table 6. Notice that the overall budget for these seven projects declined steadily from \$71.2 million to \$25.1 million, and the resulting revenue declined from \$605.1 million to \$141.9 million. If the budget had been allocated across these projects so as to maximize direct revenue, taking into account the nonlinear relationship between revenue and cost, one would expect that both the average and marginal revenue/cost ratios would increase as the budget decreased, but this did not happen.

There are undoubtedly many reasons for these observations, including the following (not necessarily listed in the order of their impact):

- In reality, the IRS does not allocate its resources solely to maximize direct revenue. Other objectives—such as maximizing the dollars assessed (rather than collected), minimizing the no-change rate, minimizing the time an audit is open, stabilizing mail volumes, and maximizing the presumed impact of the audits on the voluntary compliance of the general population—are routinely pursued, as well. Hence, the historical data upon which our curves are based are not always consistent with the revenue-maximizing assumptions we make. There are other factors that influence return selection for correspondence audits. For example, the First-Time Homebuyer Credit took effect during this period, causing correspondence audit resources to be diverted from regular projects for a couple of years to monitor this new compliance challenge.

¹⁷ Similar variations were exhibited by the other projects in this study over this time period.

- Taxpayers' compliance behavior undoubtedly changes over time. This may be particularly true on any specific tax return line item. Since correspondence audits focus on a small number of related return line items (often just one line item), the results of these audits tend to be very sensitive to trends in taxpayer behavior.

TABLE 6. Change in Revenue and Cost: "Optimal" Allocation vs. Actual Allocation Among Selected Correspondence Audit Projects, Tax Years 2006–2010 (\$ Millions)

Project	TY2006	TY2007	TY2008	TY2009	TY2010	Total
Change in Allocation of Budget						
C1	\$2.9	-\$5.3	\$3.3	\$2.0	-\$0.6	\$2.3
A1-Low	\$17.1	\$0.1	\$0.5	\$5.3	-\$0.8	\$22.3
A1-High	-\$0.8	-\$1.2	-\$0.1	-\$0.6	-\$0.2	-\$2.8
A2	-\$1.4	-\$3.6	\$0.1	\$0.0	-\$0.1	-\$5.0
A3	-\$4.1	-\$1.9	-\$1.5	-\$1.3	\$2.8	-\$5.9
A4	\$1.9	\$26.2	\$3.0	\$3.0	\$0.6	\$34.7
O	-\$15.8	-\$14.3	-\$5.3	-\$8.5	-\$1.6	-\$45.5
Total	\$0.0	\$0.0	\$0.0	\$0.0	\$0.0	\$0.0
Change in Revenue						
C1	\$20.1	-\$33.3	\$16.5	\$13.0	-\$2.8	\$13.5
A1-Low	\$141.4	\$2.3	\$5.8	\$32.7	\$0.4	\$182.6
A1-High	-\$4.3	-\$7.2	-\$0.9	-\$1.4	-\$0.5	-\$14.3
A2	-\$8.4	-\$24.8	\$0.3	\$0.3	-\$0.5	-\$33.1
A3	-\$20.3	-\$8.3	-\$6.4	-\$5.1	\$12.0	-\$28.0
A4	\$16.3	\$234.3	\$17.0	\$20.2	\$2.3	\$290.2
O	-\$81.4	-\$90.1	-\$12.6	-\$33.7	-\$3.8	-\$221.5
Total	\$63.5	\$72.8	\$19.9	\$26.1	\$7.1	\$189.4

TABLE 7. Change in Average and "Optimal" Revenue/Cost Ratios Over Time, Selected Correspondence Audit Projects, Tax Years 2006–2010

	TY2006	TY2007	TY2008	TY2009	TY2010
Total Cost (\$ Millions)	\$71.2	\$62.3	\$40.4	\$37.0	\$25.1
Total Original Revenue (\$ Millions)	\$605.1	\$543.2	\$307.9	\$273.2	\$141.9
Overall Original Average R/C	8.50	8.71	7.61	7.39	5.65
"Optimal" MR/MC ratio	6.51	8.43	3.69	5.39	4.03

6. Conclusion

This simple exercise involving just seven projects within the correspondence audit function illustrates the great potential for using this framework to increase net direct revenue through a reallocation of resources. Although the optimal allocation would be much more complicated to determine once we have estimated similar marginal revenue/cost functions for all other correspondence audit projects—and particularly for other enforcement programs—the basic approach would be the same.

In addition to paving the way for estimating similar marginal curves for other enforcement activities, this research is already leading to improved risk assessment formulas for prioritizing returns to audit. In the context of shrinking budgets and demands for improved revenue collection, this research promises to help the IRS move closer to the optimal allocation of its resources. However, the volatility of the estimated curves over time makes it very challenging to have confidence that using historical audit data to allocate resources in the current year will guide us close to the "optimal" allocation. Further research is under way to tackle this challenge.

IRS Collectibility Curve

Tom Beers, Carol Hatch, Joe Saldana, and Jeff Wilson,
Taxpayer Advocate Service Research, Internal Revenue Service

Introduction

When taxpayers incur delinquent tax liabilities, the Internal Revenue Service (IRS) sends them a series of notices during a 6-month period during which the taxpayers are in “notice status.” If the taxpayer does not resolve his or her liability during notice status, the account enters into taxpayer delinquent account (TDA) status. The IRS then determines whether the case will be referred to the Automated Collection System (ACS), assigned directly to the Collection Field function (CFf) for in-person contact by a revenue officer, assigned to the Collection Queue (“Queue”) to await assignment to a revenue officer, or shelved.¹

The ACS is a computerized inventory system and telephone call center. After a case arrives in ACS, the IRS checks for levy sources, telephone numbers, and other characteristics. These actions result in additional computer-generated notices to taxpayers. Customer Service Representatives (referred to as “Collection Representatives”²) work ACS cases and primarily respond to phone calls from taxpayers who call in response to IRS enforcement actions (*e.g.*, levies or liens) rather than proactively contacting taxpayers.

The Queue is an electronic holding bin that holds TDA accounts awaiting assignment to field revenue officers based on inventory levels.³ Cases assigned to the Queue are prioritized using a risk scoring algorithm. Shelved cases are not actively worked by the IRS while in shelved status, but continue to accumulate penalties and interest. This study does not specifically explore collections on shelved cases.

TAS was interested in examining what happens over the life of a tax debt: do people pay more of the tax debt if collections are made earlier in the debt cycle (closer to when the debt actually occurs)? Are there patterns that indicate the likelihood of collecting a debt over time? To this end, TAS Research examined the Individual Master File (IMF) Accounts Receivable Dollar Inventory (ARDI) to determine how dollars collected fluctuate as time elapses.

We looked at delinquencies that originated in each of 10 years (2003 through 2012) and analyzed those delinquencies over two time periods: the next 3 years and the next 10 years.⁴ For purposes of brevity, the tables in the body of this paper include only newly assigned TDAs in 2003, 2005, 2007, 2009, and 2011; however the Appendix contains data on TDAs newly assigned from 2003 through 2012.

Budgetary constraints will make the efficient collection of delinquencies paramount. The IRS should use data on the practical delinquency collection “window” to form the basis for its Collection policies. Good information on the time available to collect various delinquencies effectively will assist the IRS in determining what liabilities should be collected first and if it makes sense to defer the collection of smaller more current liabilities in favor of older, larger liabilities. Furthermore, this research may provide significant insights into which delinquencies the IRS should place in the Collection TDA queue and which it should shelve.⁵

Background

In past Annual Reports to Congress, the National Taxpayer Advocate noted that many of the TDAs in the IRS Automated Collection Branch and the CFf are delinquencies that have existed for several years. The following statistics highlight the age of the IRS TDA inventory:^{6,7}

¹ Shelving refers to the IRS reporting a liability as currently not collectible because of its small balance due.

² IRM 21.1.1.6.

³ Work also goes into the Queue from ACS if it cannot be resolved while in ACS status.

⁴ We chose the 10-year period for analysis because the IRS’s authority to collect delinquent taxes, *i.e.*, the collection statute, expires 10 years after the date of assessment.

⁵ The IRS places TDAs in the collection queue until a revenue officer is available to work the case.

⁶ A TDA represents only one module, generally a tax return for a single tax year. A taxpayer may have multiple TDA delinquencies.

⁷ IRS Collection Activity Report 5000-2 (Oct. 3, 2014).

- Overall, 53 percent of the IRS IMF TDA inventory has been in the function assigned the delinquency for at least 10 months (the delinquency may have been in TDA status much longer);
- Over 70 percent of the IMF TDAs in IRS inventory at the end of 2014 are Tax Year 2010 and prior liabilities; and
- Over 20 percent of the IMF TDAs have less than 4 years remaining on the collection statute, meaning that the delinquency has existed for over 6 years.

Objectives

We identified nine objectives to explore the relationship between the age of a TDA and the dollars that the IRS collects on these liabilities. These objectives explore the dollars collected as TDAs age, and differentiate between dollars collected from subsequent payments and dollars collected by offset.^{8,9} We also explore subsequent payments and offsets by various categories of the balance due amount, the type of assessment, and the accumulation of penalties and interest. Specifically, for IMF liabilities reaching TDA status, we:

- Determine amounts collected from subsequent payments on delinquencies for the 3 years after the liability reaches TDA status;
- Quantify the dollars from subsequent payments collected during the entire 10-year collection statute;
- Delineate the dollars collected from offsets of other overpayments and compare them to collections from other subsequent payments;
- Determine how the collection of liabilities varies by the amount of the delinquency;
- Determine if the rate of collection varies between self-reported liabilities and additional assessments;
- Quantify how penalty and interest cause the liability from a tax assessment to increase the total balance due;
- Determine the percent of liabilities abated by the IRS and if the percentage abated varies by the source of assessment;
- Examine the percent of cases resolved during the 10-year collection statute; and
- Determine if the percent of TDA dollars collected varies by Collection channel.

Methodology

TAS Research examined the IMF ARDI to determine how dollars collected fluctuate as time elapses. We looked at delinquencies that entered TDA status from 2003 through 2012. We analyzed liabilities entering TDA status in 2003, 2004, and 2005 for 10 years.¹⁰ We analyzed the later years through 2014. We focused initially on payments received during the first 3 years after the accounts entered TDA status. To examine payments over the 10-year collection statute and to better differentiate between subsequent payments and offsets from other taxpayer overpayments, we used transaction code data from the IMF. This allowed us to distinguish between payments and offsets, as well as to quantify abatements.¹¹ Transaction codes were also used to classify assessed interest and penalties.¹² We classified a liability by the first calendar year when it reached TDA status. If a delinquent module left and returned to TDA status, we continued to classify it by the first year the IRS assigned the liability to TDA status.¹³

We used the major source of assessment (from the ARDI file) to classify the source of assessment. Sometimes, a liability is comprised of more than one type of assessment. For example, a liability might be comprised of a self-reported assessment and an audit assessment. In this case, the type of assessment is the one most significantly contributing to

⁸ Subsequent payments include voluntary payments from taxpayers, such as those from installment agreements, and involuntary payments such as from an IRS levy.

⁹ Dollars collected from refunds or overpayments due to the taxpayer.

¹⁰ TDAs originating in 2005 will have been in notice status for several prior months. Therefore, the 10-year statute will have expired or be about to expire in 2014.

¹¹ Payments include one of the following transaction codes: 610, 611, 612, 640, 641, 642, 660, 661, 662, 666, 667, 670, 671, 672, 673, 680, 681, 682, 683, 690, 691, 692, 693, 694, 695, 760, 762, and 763. Offsets include one of the following transaction codes: 700, 701, 702, 703, 706, 710, 712, 713, 716, 720, 721, 722, 723, 730, 736, 740, and 742. Abatements include one of the following transaction codes: 161, 167, 171, 177, 181, 187, 191, 197, 235, 239, 241, 247, 271, 277, 281, 287, 291, 295, 299, 301, 305, 309, 321, 337, 341, 342, 351, 361, 538, and 549.

¹² Interest includes the following transaction codes: 190 and 196. Penalties include the following transaction codes: 160, 166, 170, 176, 270, 276, 280, 286, 320, and 350.

¹³ A delinquent account can leave TDA status and enter into another status. For example, if the taxpayer enters into an installment agreement (IA) to repay the delinquency, the account leaves TDA status and enters into IA status. If the taxpayer subsequently defaults on the IA, the account will reenter TDA status.

the balance owed. We determined whether the IRS assigned a TDA liability to ACS, collection queue, or Cff by the Taxpayer Service Returns Processing Category (TRCAT) code, which differs depending on where a liability is located in the collection stream.

Limitations

When we discuss changes in the total module balance of TDAs, we have included both assessed and accrued penalties and interest. However, in the specific objective regarding penalties and interest balance, we have tracked only assessed penalties and interest but have not quantified accrued penalties and interest. Additionally, interest assessed amounts do not contain restricted interest assessments.¹⁴ Although it is a relatively small portion of abatements, dollars abated as a result of accepted offers in compromise are included in total abatements.¹⁵

Findings

In this section, we present the findings for each of the objectives. In addition to providing the data pertinent to each objective, we also offer some insights on whether the results are changing over time and why the underlying trends are present.

Determine amounts collected from subsequent payments on delinquencies for the 3 years after the liability reaches TDA status.¹⁶

For TDAs originating after 2003, our analysis showed that: (a) dollars collected decrease by more than 50 percent from the first year to the second year; and (b) in the third year, collections decrease by about one-third from the amount collected in the second year.¹⁷ In other words, collections are over twice as much during the first year as in the following year and over three times the collection in the third year. For TDAs originating in 2007, collections declined by about 64 percent during the second year after the cases entered TDA status. For 2009, the decrease in total dollars collected in the third year was only about 27 percent. Nevertheless, overall collections for cases entering TDA status after 2003 decreased by about two-thirds from the first year to the third year after entering TDA status.

Table 1 depicts these findings by the years elapsed since the initial liability reached TDA status:

TABLE 1. Subsequent Payments (in \$ Millions) Decrease as Time Elapses, Selected Years Assigned TDA

Years Elapsed	2003		2005		2007		2009		2011	
	Subsequent Payments	% Decrease in Collections from Prior Year	Subsequent Payments	% Decrease in Collections from Prior Year	Subsequent Payments	% Decrease in Collections from Prior Year	Subsequent Payments	% Decrease in Collections from Prior Year	Subsequent Payments	% Decrease in Collections from Prior Year
1	\$1,786.4		\$2,990.8		\$3,664.8		\$3,631.9		\$3,800.1	
2	\$1,166.8	-35%	\$1,344.1	-55%	\$1,330.4	-64%	\$1,675.5	-54%	\$1,748.1	-54%
3	\$848.5	-27%	\$ 832.6	-38%	\$907.0	-32%	\$1,216.8	-27%	\$1,177.6	-33%

Despite accumulation of penalty and interest, as the IRS collects additional dollars, the balance due declines over time.¹⁸ Table 2 shows the overall decline in total module balance over the first 3 years after the liability reached TDA status:

¹⁴ Restricted interest is assessed by transaction code 340 (and abated by transaction code 341). Restricted interest arises when any portion of the interest on an overpayment or underpayment is calculated from a date other than the one that applies to the return as filed. This happens most often when there is a carryback of a loss or credit.

¹⁵ The Fiscal Year (FY) 2014 liabilities compromised were 1.2 percent of the amount of TDAs at the beginning of FY 2014.

¹⁶ Subsequent payments include voluntary payments from taxpayers such as those from IAs and involuntary payments such as from an IRS levy.

¹⁷ In 2003, collections of new TDAs decreased by only about 35 percent from the first to the second year, even though the decrease from the second to the third year was similar to later years. See the Appendix for complete details on all years studied.

¹⁸ This is true only if the dollars collected exceed penalty and interest accruals. In an earlier study examining only currently not collectible (CNC) cases, the module balance actually increased as time elapsed.

TABLE 2. Rate of Module Balance Decline Slows, Selected Years Assigned TDA

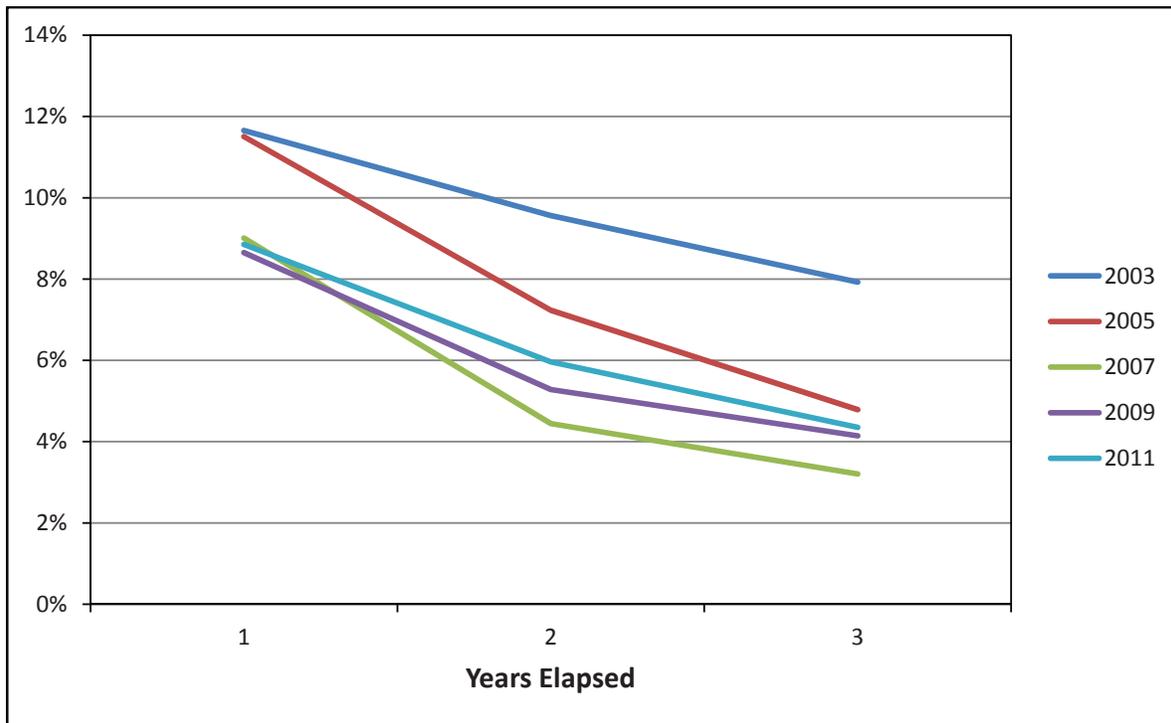
Years Elapsed	2003		2005		2007		2009		2011	
	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance
0	\$15,326.2		\$25,996.1		\$40,678.5		\$41,987.7		\$42,926.2	
1	\$12,321.3	-20%	\$20,872.6	-20%	\$32,783.3	-19%	\$35,332.5	-16%	\$34,795.8	-19%
2	\$10,370.3	-16%	\$17,657.4	-15%	\$28,948.3	-12%	\$31,581.2	-11%	\$29,792.6	-14%
3	\$8,841.3	-15%	\$15,759.1	-11%	\$26,531.7	-8%	\$28,767.3	-9%	\$27,132.4	-9%

Comparing the two previous tables, one notices that the module balance decreases more rapidly than the dollars collected would indicate. This occurs because of the complete or partial abatement of some liabilities, particularly during the first 2 years of a delinquency. We will explore abatements in greater detail in a subsequent section.

On a percentage basis, the dollars collected drop significantly from the first year to the second year, but the decrease slows in the third year. We will explore this issue further in the next study objective when we look at the entire 10-year statutory period to collect delinquent tax liabilities.

Even though the original module balance is declining, the percent collected of the balance is also declining as illustrated in Figure 1:

FIGURE 1. Decline in Dollars Collected as a Percent of Module Balance



An analysis of the data shows that dollars collected decrease as a liability ages. Dollars collected as a percentage of the prior-year dollars collected also decline significantly. Finally, the percent of the original TDA liability, including penalties and interest, being collected decreases significantly from the first year to the second year and then continues to decline, but at a slower rate. Accordingly, the rate at which the total amount of the delinquency decreases slows as time progresses.

Quantify the dollars from subsequent payments collected during the entire 10-year collection statute.

In the first objective, we looked at the first 3 years of collections after a liability reached TDA status. We looked at a period of 3 years because private collection agencies believe that nearly all monies on delinquent debts are collected within the first 3 years after the debt becomes due. Next, we will examine what happens over the entire statutory 10-year collection period.

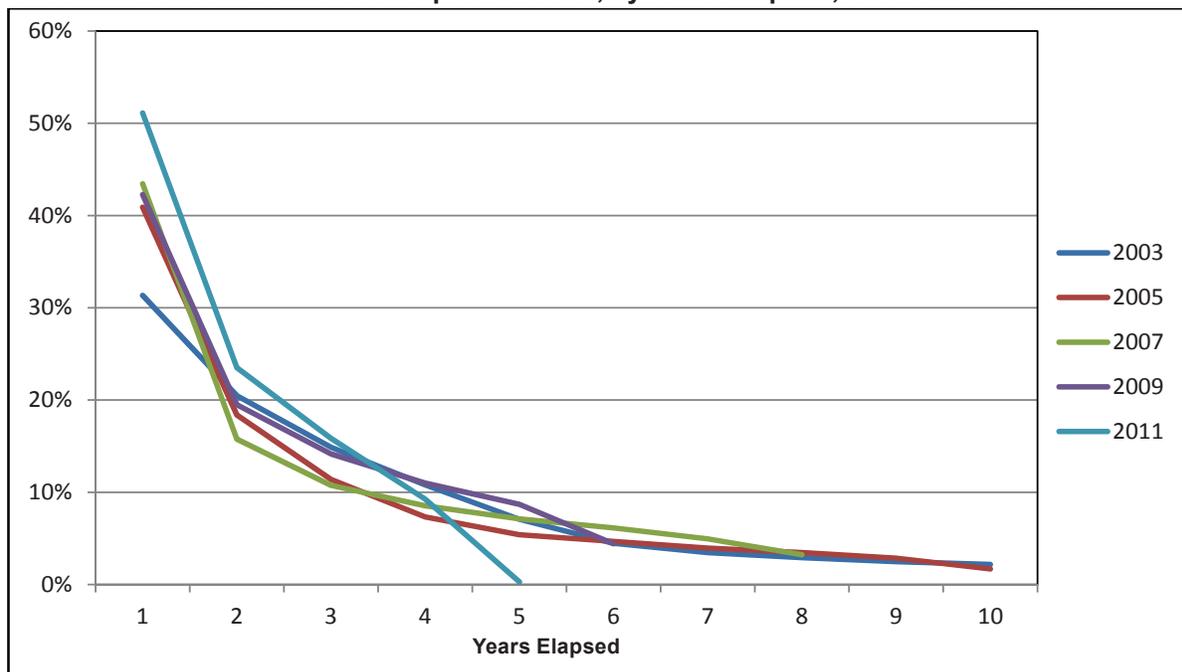
Table 3 depicts the subsequent payments by years elapsed since TDA issuance and the percent of the total dollars collected in each year:

TABLE 3. Subsequent Payments as a Percent of Total Subsequent Payments Collected Per Year, Selected Years Assigned TDA¹⁹

Years Elapsed	2003		2005		2007		2009		2011	
	Amount (\$ Millions)	% of Total								
1		31%	\$2,990.8	41%	\$3,664.8	43%	\$3,631.9	42%	\$3,800.1	51%
2		20%	\$1,344.1	18%	\$1,330.4	16%	\$1,675.5	19%	\$1,748.1	24%
3	\$848.5	15%	\$832.6	11%	\$907.0	11%	\$1,216.8	14%	\$1,177.6	16%
4	\$615.1	11%	\$535.8	7%	\$720.3	9%	\$944.8	11%	\$688.5	9%
5	\$402.9	7%	\$394.7	5%	\$600.3	7%	\$746.6	9%	\$20.8	0%
6	\$254.2	4%	\$341.3	5%	\$517.4	6%	\$379.5	4%		
7	\$196.6	3%	\$289.5	4%	\$417.4	5%				
8	\$166.0	3%	\$252.3	3%	\$272.5	3%				
9	\$141.4	2%	\$209.5	3%	\$7.8	0%				
10	\$123.3	2%	\$123.6	2%						
Total		100%	\$7,314.3	100%	\$8,437.9	100%	\$8,595.2	100%	\$7,435.1	100%

Figure 2 illustrates this same information:

FIGURE 2. Percent Collected in up to 10 Years, by Years Elapsed, Five Selected Years



¹⁹ Subsequent payments include voluntary payments from taxpayers such as those from installment agreements and involuntary payments such as from an IRS levy.

Table 3 clearly shows a decline in the dollars collected as time elapses throughout the collection statute period. Dollars collected level off at about 2 percent in the last year of the collection statute. As we saw in the first objective, the total balance due also declines, although much more slowly in the latter years. This trend is also illustrated in Table 4.

TABLE 4. Decline in Total Balance Owed Within Ten Years After TDA Origination, Selected Years Assigned TDA²⁰

Years Elapsed	2003		2005		2007		2009		2011	
	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance	Total Module Balance (\$M)	% Decline in Module Balance
0	\$15,326.2		\$25,996.1		\$40,678.5		\$41,987.7		\$42,926.2	
1	\$12,202.9	20%	\$20,955.2	19%	\$32,849.0	19%	\$34,910.1	17%	\$34,032.3	21%
2	\$10,705.9	12%	\$18,585.0	11%	\$29,935.1	9%	\$31,718.6	9%	\$29,319.0	14%
3	\$9,603.3	10%	\$17,390.0	6%	\$28,301.1	5%	\$29,367.1	7%	\$27,055.1	8%
4	\$8,947.3	7%	\$16,596.2	5%	\$26,943.5	5%	\$27,478.0	6%	\$26,304.4	3%
5	\$8,477.7	5%	\$15,982.9	4%	\$25,668.2	5%	\$26,092.4	5%		
6	\$8,148.7	4%	\$15,505.7	3%	\$24,806.1	3%	\$25,649.1	2%		
7	\$7,835.7	4%	\$15,067.6	3%	\$24,032.8	3%				
8	\$7,522.2	4%	\$14,613.4	3%	\$23,740.4	1%				
9	\$7,139.4	5%	\$14,138.7	3%						

We should note that the total module balance continues to decline because some accounts are paid in full as time progresses. However, for those accounts that are not resolved, their penalties and interest continue to rise. A larger decrease in year 10 occurs because the collection statute has ended for a majority of the liabilities, and the IRS then clears the previous balance due.

As dollars are collected, the balance due declines over time. Abatements also decrease the liabilities. However, penalties and interest increase the total amount due. We examined the amount of dollars collected by subsequent payments as a percent of the module balance at the beginning of each one-year period. Even though the total balance due generally decreases as taxpayers make subsequent payments and offsets and the IRS abates some portion of the assessment, the percent decrease also shows a similar decline in each year during the study period, as illustrated in Table 5.

TABLE 5. Year-to-Year Percent Decline in Total Balance Due, Selected Years Assigned TDA

Years Elapsed	2003	2005	2007	2009	2011
1	12%	12%	9%	9%	9%
2	10%	6%	4%	5%	5%
3	8%	4%	3%	4%	4%
4	6%	3%	3%	3%	3%
5	5%	2%	2%	3%	
6	3%	2%	2%	1%	
7	2%	2%	2%		
8	2%	2%	1%		
9	2%	1%			
10	2%	1%			

²⁰ The ending balance after 10 years is not shown. Since the 10-year collection statute generally expires in the 10th year after the IRS assigns a case to TDA status, the module balance becomes significantly reduced by the abatements of liabilities that the IRS is no longer permitted to collect.

As a percentage of the balance due, dollars collected generally drop most precipitously from the first to the second year. As the table indicates, the ratio of dollars collected to balances due drops as elapsed time increases.

Determine the dollars collected from offsets of other overpayments and compare to collections from other subsequent payments.

Analysis of the collection activity reports for a number of years shows that a significant percentage of the total dollars collected come from refund offsets, particularly in ACS. Therefore, we distinguished between dollars collected through subsequent payments²¹ and dollars collected through offsets from overpayments on other tax modules. Table 6 compares the amount and percent of the initial balance due collected by subsequent payments to collections by offsets from overpayments (refunds) on other tax accounts (generally other tax years).

TABLE 6. Dollars Collected and Offset, Selected Years (\$ in Millions)

Year Assigned TDA	Balance Due	Subsequent Payments	% Collected	Amount Offset	% Offset	Total % Collected
2003	\$15,326.2	\$5,701.2	37.2%	\$ 2,150.7	14.0%	51.2%
2005	\$25,996.1	\$7,314.3	28.1%	\$ 3,086.5	11.9%	40.0%
2007	\$40,678.5	\$8,437.9	20.7%	\$ 4,493.5	11.0%	31.8%
2009	\$41,987.7	\$8,595.2	20.5%	\$ 4,173.6	9.9%	30.4%
2011	\$42,926.2	\$7,435.1	17.3%	\$ 3,583.2	8.3%	25.7%

For delinquencies reaching TDA status in 2003, the amount collected from subsequent payments is nearly three times the amount offset. However, for delinquencies reaching TDA status in later years, subsequent payments are only about twice the amount offset. On a percentage basis to the amount initially owed, subsequent payments have decreased significantly from TDAs first assigned in 2003 to TDAs first assigned in 2011; however, offsets have remained relatively stable, decreasing by only a few percent. While it is true that delinquencies reaching TDA status since 2006 still have some years remaining on the collection statute, the dollars collected increased by less than 10 percent during the last 6 years of the 10-year collection statute for TDAs issued in 2003 and 2005. Therefore, it is unlikely that dollars collected from TDAs issued in later years will increase sufficiently to realize the same proportion of dollars collected to dollars offset as in earlier years. Since offsets are relatively flat over the period examined, we generally see the same trends in total dollars collected, as we saw when examining only subsequent payments.

Determine how the collection of liabilities varies by the amount of the delinquency.

In addition to comparing the dollars collected by subsequent payments and the offsets of overpayments, we also compare the dollars collected by subsequent payments and offsets in six ranges of the balance due. As one might expect, the IRS collects a greater percentage of the liability when it is not more than \$5,000.

As illustrated in Table 7, an analysis of the TDA modules clearly shows that the majority of delinquency amounts do not exceed \$5,000. However, higher dollar ranges contain the highest percentage of the delinquent dollars, even though these categories contain only a small percent of the delinquent modules. For example, in 2003, about three-quarters of the TDA modules were under \$5,000, while over 80 percent of the delinquent dollars were in the highest three balance due ranges, *i.e.*, the categories greater than \$5,000. In fact, over half of the overall delinquent dollars were on modules with more than \$25,000 due. Interestingly, however, from 2003 to 2011, the percent of delinquent TDA modules under \$5,000 fell from over 75 percent to under 68 percent while the percent of dollars in the highest three dollar ranges increased from under 82 percent to over 88 percent. This trend indicates that more taxpayers owe liabilities over \$5,000.²² Inflation undoubtedly accounts for part of this increase, rising by about 17 percent during this period, but the combined initial TDA balance for modules with balances greater than \$5,000 is nearly three times as high in 2011 as in 2003.²³ This increase in balance due is a disturbing trend for the IRS.

²¹ Subsequent payments include voluntary payments from taxpayers such as those from installment agreements and involuntary payments such as from an IRS levy.
²² For liabilities entering TDA status in 2009, only about 60 percent of the delinquent modules had liabilities of \$5,000 or less. This situation may be attributable to the depressed economic conditions in 2008.
²³ Bureau of Labor Statistics Consumer Price Index inflation calculator available at: http://www.bls.gov/data/inflation_calculator.htm.

TABLE 7. Modules, Balance Due, and Dollars Collected by Initial Module Liability Dollar Range

Year	Description	\$1 to \$1,000	\$1,001 to \$2,000	\$2,001 to \$5,000	\$5,001 to \$10,000	\$10,001 to \$25,000	Greater Than \$25,000
2003	Module Count	451,712	505,146	565,164	250,331	160,431	92,971
	Percent of Modules in Range	22%	25%	28%	12%	8%	5%
	Aggregate Balance Due	\$240.8	\$740.6	\$1,793.1	\$1,745.4	\$2,446.9	\$8,359.3
	Percent of Total Balance in Range	2%	5%	12%	11%	16%	55%
	Percent Collected by Subsequent Payment	66%	49%	49%	48%	43%	29%
	Percent Collected by Offset	50%	44%	35%	22%	13%	5%
2005	Module Count	467,988	561,662	762,610	388,628	254,399	172,255
	Percent of Modules in Range	18%	22%	29%	15%	10%	7%
	Aggregate Balance Due	\$250.8	\$832.9	\$2,462.5	\$2,713.9	\$3,886.6	\$15,849.4
	Percent of Total Balance in Range	1%	3%	9%	10%	15%	61%
	Percent Collected by Subsequent Payment	79%	54%	50%	44%	37%	18%
	Percent Collected by Offset	55%	50%	38%	23%	13%	3%
2007	Module Count	781,534	666,064	1,006,717	616,892	408,744	260,839
	Percent of Modules in Range	21%	18%	27%	16%	11%	7%
	Aggregate Balance Due	\$449.3	\$978.5	\$3,313.1	\$4,309.5	\$6,214.7	\$25,413.3
	Percent of Total Balance in Range	1%	2%	8%	11%	15%	62%
	Percent Collected by Subsequent Payment	60%	45%	40%	33%	27%	13%
	Percent Collected by Offset	61%	51%	37%	23%	12%	3%
2009	Module Count	520,936	596,584	1,038,155	697,679	479,893	292,604
	Percent of Modules in Range	14%	16%	29%	19%	13%	8%
	Aggregate Balance Due	\$290.8	\$907.6	\$3,388.5	\$4,874.9	\$7,346.0	\$25,179.9
	Percent of Total Balance in Range	1%	2%	8%	12%	17%	60%
	Percent Collected by Subsequent Payment	58%	40%	35%	27%	23%	15%
	Percent Collected by Offset	46%	40%	31%	19%	10%	4%
2011	Module Count	825,154	754,679	1,136,688	639,600	422,102	246,137
	Percent of Modules in Range	20.5%	18.8%	28.2%	16%	10%	6%
	Aggregate Balance Due	\$480.4	\$1,117.9	\$3,718.7	\$4,484.3	\$6,436.7	\$26,688.3
	Percent of Total Balance in Range	1%	3%	9%	10%	15%	62%
	Percent Collected by Subsequent Payment	37%	27%	22%	18%	17%	11%
	Percent Collected by Offset	47%	39%	27%	16%	9%	2%

We also see in Table 7 that the percent of dollars offset is highest in the lowest dollar categories of TDA dollars due, declining as the balance due increases. As the table indicates, about half of delinquency amounts up to \$2,000 are collected by refund offsets. Since a majority of the TDAs in ACS have lower balances due, it is not surprising that almost half of the ACS total dollars collected are from offsets.²⁴ The dollars collected from offsets also decline as the TDA balance due increases.

We see from Table 7 that more than 100 percent of the initial balance is sometimes paid. This occurs because penalties and interest have continued to accrue so the final balance paid by the taxpayer is significantly higher than the initial balance due.

Determine if the rate of collection varies between self-reported liabilities and additional assessments.

We explored whether the amount collected by the IRS depends on the source of the underlying assessment. Specifically, we examined whether the IRS collects a greater percentage of self-reported liabilities than liabilities initiated or increased by the IRS (e.g., additional assessments from audits, third-party information matching (AUR), or Automated

²⁴ Collection Activity Report 5000-2 (Oct. 2014). For individual liabilities, offsets actually exceeded dollars collected through collection activities and voluntary subsequent payments.

Substitute for Returns). As expected, the IRS is more successful at collecting self-reported liabilities than additional assessments. Table 8 depicts the difference between percentages of the initial liability collected by subsequent payment, based on the source of the liability.

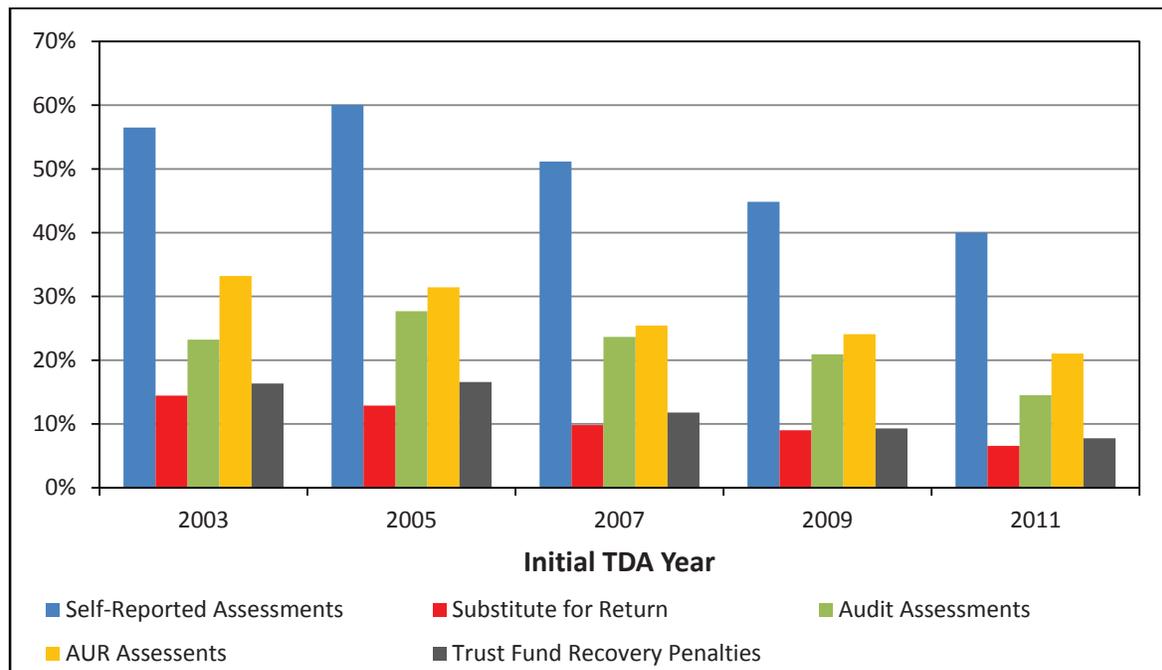
TABLE 8. Percent Collected by Subsequent Payment Based on Source of Assessment

Year	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2003	56%	14%	23%	33%	16%
2005	60%	13%	28%	31%	17%
2007	51%	10%	24%	25%	12%
2009	45%	9%	21%	24%	9%
2011	40%	7%	15%	21%	8%

Clearly, the IRS is most likely to collect self-reported liabilities, which it does at a rate at least twice as great as it collects audit assessments.²⁵ In general, the IRS collects a slightly higher percentage of AUR assessments than audit assessments. The IRS also collects only a small percentage of substitute for returns and trust fund recovery penalty assessments.

Figure 3 illustrates the difference in the percent of the initial liability collected by subsequent payment for various assessment types.

FIGURE 3. Percent of Initial TDA Liability Collected by Subsequent Payment, Based on Assessment Type



Interestingly, the dollars collected on all of these types except audits have declined significantly since 2005. This disturbing trend merits additional investigation.²⁶

²⁵ It seems reasonable that taxpayers who assess themselves a balance due are more willing to pay than those who are audited. This may also reflect the fact that returns expected to generate larger audit assessments tend to be selected for audit and, as our analysis shows, a smaller percent of large liabilities—*i.e.*, liabilities exceeding \$5,000—are ultimately collected.

²⁶ Since the collection statute has not expired for cases reaching TDA status in the latter years shown in the chart, more monies will be collected; however, as we have shown, we do not expect the IRS to collect many more dollars on these liabilities in the last half of the collection statutory period.

We also broke out offsets from the total dollars collected and explored the dollars collected due to offsets. The IRS collects a higher percentage of AUR assessments through offsets than any other type of assessment, even self-reported assessments. Table 9 displays the percent of the initial TDA balance offset by source of assessment.

TABLE 9. Percent Collected by Offsets Based on Source of Assessment

Year	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2003	18%	4%	12%	34%	6%
2005	20%	5%	20%	32%	6%
2007	20%	5%	25%	36%	6%
2009	15%	4%	20%	28%	6%
2011	10%	2%	12%	25%	4%

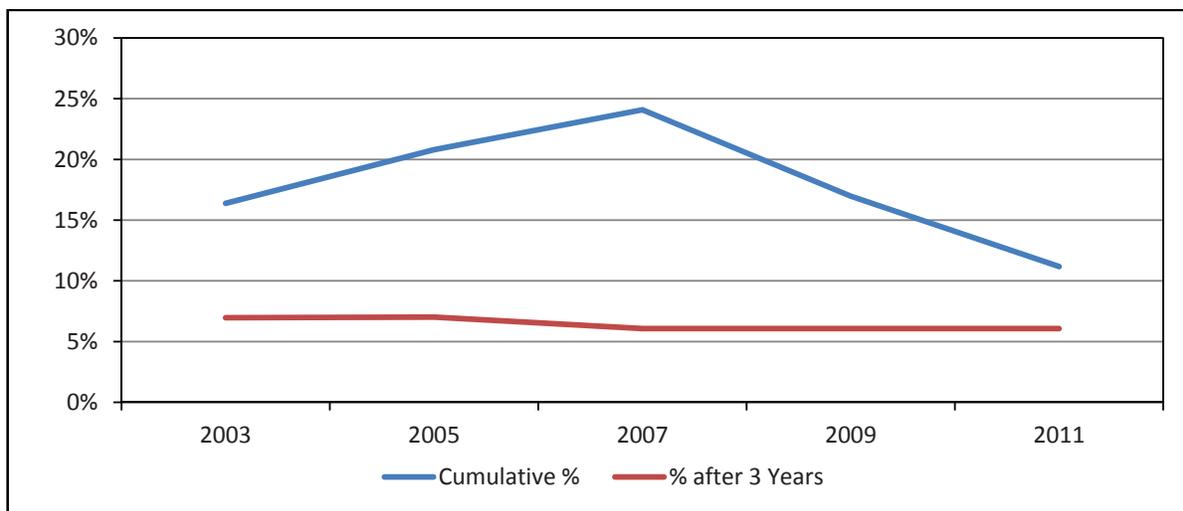
Also, the difference in offset dollars collected between audit and self-reported assessments is not as great as the difference of offset dollars collected between audit and AUR additional assessments. In fact, AUR assessments actually resulted in the highest percent of the liabilities paid by offset—almost twice that of self-reported liabilities.

AUR liabilities also account for three times the percent of dollars offset to audit liabilities in 2003. While the gap in dollars offset between AUR and audit liabilities has narrowed by 2011, it is still significant. Perhaps the reason AUR assessments see such a high percent of their initial TDA balance offset, even compared to self-reported liabilities, is because a much higher percent of self-reported liabilities are collected through subsequent payments. Taxpayers who receive AUR assessments may also be more likely to receive future refunds, since these taxpayers are often wage earners who have their income tax withheld by the payer.

Quantify how penalty and interest cause the liability from a tax assessment to increase the total balance due.

At first glance, it appears that penalties and interest have been declining since 2003. However, the significant abatement rate of the initial liability masks the increase in the balance due attributable to penalties and interest. Specifically, abatements have increased so the original TDA balance has experienced a greater decrease. Therefore, penalties and interest comprise a greater percentage of the amount actually determined due by the IRS. When one considers the amounts of abatement from the initial TDA assessment, the percentage of the liability actually due to penalties and interest is generally rising. From 2003 to 2007, penalties have comprised a larger portion of the initial TDA balance the IRS has assessed and determined due. For 2009 and 2011, sufficient time has not elapsed to realize the significant effect of penalties and interest. Figure 4 illustrates this fact, showing that through 2007, penalties and interest have continued to constitute a larger percentage of the initial liability the IRS has determined due.

FIGURE 4. Percent of Liability Due Attributable to Penalties and Interest



From 2003 to 2007, the portion of the initial assessment (actually due) resulting from penalties and interest increased to almost a quarter of the initial liability. For later years, the portion on the initial TDA liability (actually due) is likely to become an even higher percentage, although sufficient time has not elapsed to experience the full impact of penalties and interest.

As the IRS takes longer to collect liabilities, taxpayer burden will continue to increase, as taxpayers pay even larger amounts of penalties and interest. The graph also shows that through the first 3 years after TDA assignment, penalties and interest remain relatively constant. However, as the IRS continues to resolve fewer TDAs, the percent of the initial liability attributable to penalties and interest will continue to grow. By the 10th year of the collection statute, taxpayers with TDAs originating in 2003 and 2004 owed more than twice the amount of penalties and interest they owed 3 years after TDA assignment. For TDAs originating in 2005, taxpayers owed more than three times the penalty and interest in 2014 (10 years later) than they did in 3 years after the initial TDA. As the years progress, the IRS has assigned more accounts to TDA status; however, for the 3 years the 10-year collection statute has had sufficient time to lapse, the average amount of assessed penalty and interest has also increased for each TDA.

Determine the percent of liabilities abated by the IRS and if the percentage abated varies by the source of assessment.

Both dollars abated from the initial TDA assessment²⁷ and the percent of the initial balance abated have continued to be higher than the 2003 rate, and they remain at an overall higher level, as indicated in Table 10.

TABLE 10. Percent of Initial TDA Balance Abated²⁸

Year	Initial TDA Balance	Amount Abated	Percent Abated
2003	\$15,326,191,192	\$2,985,977,270	19%
2005	\$25,996,084,845	\$8,066,761,341	31%
2007	\$40,678,451,308	\$13,086,103,480	32%
2009	\$41,987,700,518	\$10,716,623,485	26%
2011	\$42,926,217,917	\$11,990,870,525	28%

The dollars abated continue to increase. The rate of abatement for 2007 is higher than in 2003 and 2005, even though the TDAs in 2007 have about two more years remaining on the collection statute. The abatement rate is down slightly since 2007; however, less time has elapsed. The data suggest that Collection is continuing to focus significant resources on bad assessments.

We also explored the TDA dollars abated by the source of assessment, as indicated in Table 11. IRS substitute for return assessments are the most likely to be abated.²⁹ For 2003 and 2005, where 10 years have elapsed since assignment to a TDA, almost half of liability amounts have been abated.

TABLE 11. Percent of TDA Amount Abated, by Source of Assessment

Year	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2003	6%	49%	15%	15%	39%
2005	6%	47%	12%	29%	40%
2007	12%	43%	14%	28%	35%
2009	9%	36%	13%	27%	28%
2011	16%	40%	19%	18%	29%

²⁷ Dollars abated may include tax, penalty, and interest.

²⁸ For TDAs initially assigned in 2003 and 2005, abatements are also attributable to the expiration of the 10-year collection statute.

²⁹ This is presumably due to the fact that SFR assessments are based on the assumption that the taxpayer is single, claiming the standard deduction. That assessment prompts some taxpayers to file a delinquent return, which documents a lower tax liability—thus, the abatement of the overstated liability. However, the remaining assessment could still be very cost-effective to collect.

Obviously, substitute for return assessments are generating considerable rework for the IRS and may be preventing the IRS from collecting additional subsequent payments on more productive work. IRS should ensure substitute for return assessments are at least as cost-effective as other types of assessments and review current procedures to identify revisions that could improve productivity.

The abatement rate of AUR assessments has also increased significantly since 2003, possibly implying that the IRS is selecting more cases for AUR assessments, even though it is less certain that the taxpayer is liable for the additional tax. Trust Fund Recovery Penalties (TFRP) have an abatement rate almost as high as that of substitute for return assessments. However, TFRP assessments may have necessarily high abatement rates because the IRS abates the liability, as it is paid by the underlying corporation or other responsible officers.

Examine the percentage of cases resolved during the 10-year collection statute.

We examined the percentage of cases completely resolved within the usual 10-year collection statute. Overall, the IRS completely resolved nearly 80 percent of the cases reaching TDA status in 2003 and 2005 by the ninth year of the collection statute.³⁰ The percentage of cases closed in the 10th year of collection statute increases significantly because liabilities are being abated in full as the collection statute expires.³¹ Although more time remains on the collection statute for TDAs assigned in more recent years through equivalent periods of elapsed time, the percent of the balance due collected has been declining from earlier years.³² This information is illustrated by Table 12.

TABLE 12. Cumulative Closure Rate

Elapsed Years	2003	2005	2007	2009	2011
1	23%	25%	27%	21%	25%
2	38%	39%	40%	33%	37%
3	50%	49%	47%	41%	45%
4	58%	56%	53%	48%	48%
5	65%	61%	58%	52%	
6	69%	65%	62%	54%	
7	73%	68%	65%		
8	76%	71%	66%		
9	80%	74%			
10	95%	80%			

Though the IRS resolves most TDA modules, at least one-third of the total dollar amount of the liabilities remains 4 years after a delinquency reaches TDA status, as illustrated in Figure 5.³³

The overall high closure rate is undoubtedly because, as discussed earlier, the vast majority of modules owe no more than \$5,000. The IRS is generally effective at collecting these smaller liabilities through subsequent payments and offsets. The data also indicate that the percentage of the total liability collected, including penalties and interest, has been declining since 2003, although the rate of liability growth due to penalties and interest has increased.

As the closure rate has generally declined from 2003 to 2009, the volume of TDA cases remaining open has continued to increase. Table 13 shows the volume of cases still open since the liability was assigned to TDA status.

³⁰ The liability may be completely resolved because: (a) the taxpayer paid the liability in full, including penalties and interest; (b) the IRS may have determined the liability was incorrect and abated all or part of it; or (c) the IRS may have accepted an offer to compromise the tax liability for less than the full amount.

³¹ The closure rates depicted are for TDA liabilities. Since time has elapsed between the assessment of a liability and when the IRS assigns it to TDA status, the collection statute generally expires during the 10th year since the liability reached TDA status (rather than at the end of the 10th year). Certain actions, such as the consideration of an installment agreement, offer in compromise, or bankruptcy proceeding may extend the collection statute. Additionally, the taxpayer may voluntarily extend the collection statute, usually to pursue a long-term installment agreement.

³² The closure rate for 2011 is higher than the rate in 2009 until the fourth year.

³³ We used the fourth year of the collection statute for an even comparison.

FIGURE 5. Liability Remaining Four Years After TDA Issued

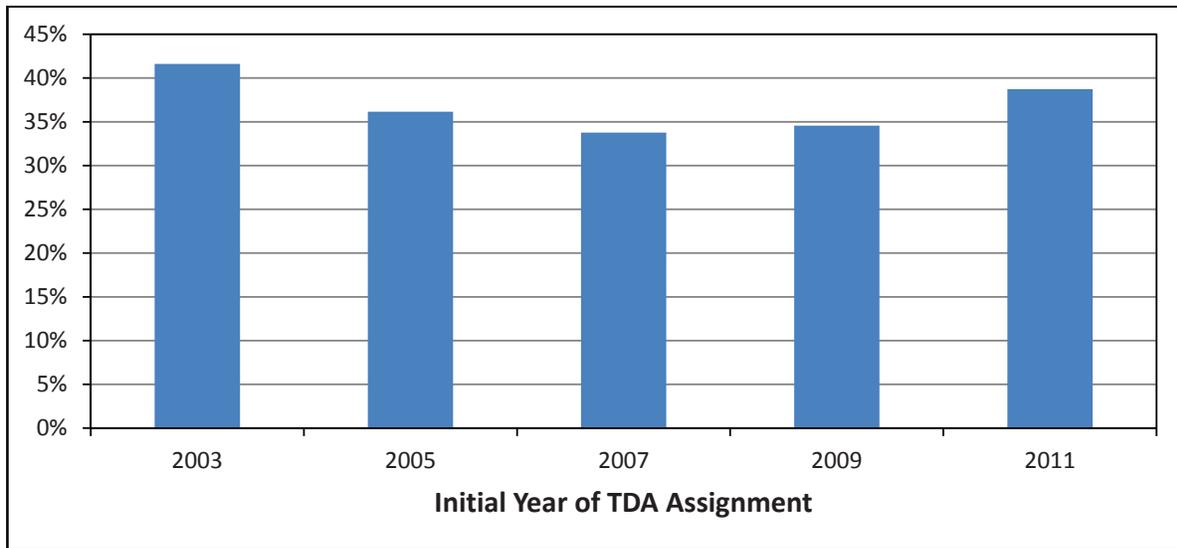


TABLE 13. Percent of Cases Remaining Open by Years Since Becoming a TDA

Year Since TDA Issued	2003	2005	2007	2009	2011
1	77%	75%	73%	79%	75%
2	62%	61%	60%	67%	63%
3	50%	51%	53%	59%	55%
4	42%	44%	47%	52%	52%
5	35%	39%	42%	48%	
6	31%	35%	38%	46%	
7	27%	32%	35%		
8	24%	29%	34%		
9	20%	26%			
10	5%	20%			

The volume of open cases in 2011 is many times larger than in 2003. A significant reason for this is that the volume of new TDAs has increased so dramatically; another might be the declining trend of Collection staffing. The table demonstrates that the closure rate drops as the years progress after a module reaches TDA status. While one-fifth or less of the cases remained unresolved at the time of collection statute expiration for new TDAs from 2003 to 2005, it is likely that nearly a third of the new TDAs since 2007 will remain unresolved at the time of collection statute expiration.

Determine if the percentage of TDA dollars collected varies by Collection channel

The dollars collected and abated do vary by Collection channel. Table 14 shows that the largest percentage of dollars collected by subsequent payments and refund offsets were garnered by ACS.

TABLE 14. Percent of Initial Balance Satisfied by Payments, Offsets, or Abatements

Year	ACS			Queue			Collection Field Function		
	Percent of Initial Balance			Percent of Initial Balance			Percent of Initial Balance		
	Collected by Subsequent Payments	Collected by Offsets	Abated	Collected by Subsequent Payments	Collected by Offsets	Abated	Collected by Subsequent Payments	Collected by Offsets	Abated
2003	44%	22%	14%	29%	6%	23%	32%	6%	28%
2005	40%	21%	21%	17%	5%	37%	21%	6%	39%
2007	39%	25%	19%	11%	4%	36%	13%	5%	41%
2009	30%	16%	20%	9%	4%	27%	13%	5%	32%
2011	26%	15%	16%	8%	3%	37%	12%	2%	42%

The table shows that ACS dollars collected from subsequent payments have continued to decrease since 2003. For liabilities reaching TDA status since 2005, additional time remains to receive subsequent payments and offsets; however, the percent of the liability collected has increased by no more than 10 percent in the final 6 years of the collection statute. Therefore, it seems likely ACS will collect a significantly smaller percentage of the initial TDA balance than in 2003. The trend of the IRS collecting fewer dollars through subsequent payments is even stronger for the cases assigned to the queue and Cff.

Offsets as a percentage of the initial TDA balance due actually increased slightly for new ACS TDAs from 2003 to 2007, but then drastically decreased in 2009 and 2011. For TDAs assigned to the queue or Cff, offsets as a percent of the initial TDA balance have generally remained constant, though garnering a relatively small percent of the initial TDA balance.

Abatements of at least some of the initial TDA balance are relatively high in all three functions with TDA inventory. However, the percentage of the initial TDA balance abated is higher in the queue than in ACS and even higher in Cff. In fact, about a third of the initial balances of the TDAs assigned to Cff are abated. This means that Cff personnel are spending a significant portion of their time resolving problem assessments. Accordingly, a review of current procedures might identify ways that these cases could be worked more effectively.

After removing abatements from the initial balance due and when considering only the first 6 years since the case reached TDA status, the percent of initial TDA dollars collected is significantly higher, as indicated by Table 15.³⁴

TABLE 15. Percent of Initial TDA Balance After Abatements Collected by Payments and Offsets After First Six Years in TDA Status

Year	ACS	Queue	Cff
2003	67%	39%	45%
2005	69%	28%	35%
2007	73%	20%	26%
2009	58%	19%	26%

Although Table 15 combines dollars collected through subsequent payments and offsets, the total amount collected becomes a larger percent of the actual balance due, since abatements are excluded. This is particularly noticeable in Cff, which consistently has the highest percentage of abatements when compared to the other TDA collection channels. In general, the percent of the initial TDA balance collected has declined since 2003.

Conclusions and Summary

The IRS is more successful at collecting liabilities soon after TDA assignment. This result is similar to the experience of private collection agencies. Dollars do continue to be collected throughout the life of the 10-year collection statute period; however, the payment rate slows significantly. As one might expect, the IRS is also more successful in its

³⁴ We have removed TDAs originating in 2011 since sufficient time has not elapsed to examine collections 6 years later.

collection of self-reported assessments and smaller TDA balances. The IRS continues to deal with a high number of bad assessments that hamper its TDA collections. While we are heartened by the IRS's willingness to abate improper (or uncollectible) assessments, we wonder how many taxpayers pay assessments for which they are not liable, before the IRS even assigns the delinquency to TDA status. We have distilled the findings from the nine objectives into nine specific conclusions.

1. Dollars collected in aggregate and as a percentage of the balance due decrease significantly during the first 3 years after the IRS assigns a liability to TDA status. The decline in the module balance also slows significantly during these first 3 years.
2. When continuing to look at the collection of liabilities after the third year of the initial TDA assignment, collections continue to dwindle, and the reduction in the module balance declines almost completely by the expiration of the collection statute.
3. Overall, dollars collected through the offsets of other overpayments are significantly less than dollars collected through subsequent payments. However, dollars collected through offsets decrease much less precipitously than dollars collected from subsequent payments as time elapses from the initial TDA assignment.
4. Delinquent modules with balances due not in excess of \$5,000 comprise the vast majority of TDAs. However, over 80 percent of the total amount due resides with TDAs with balances greater than \$5,000. The IRS collects both a higher percentage of subsequent payments and offsets in the lowest balance due categories. Collections and offsets as a percentage of the balance due progressively decrease as the balance due rises.
5. The percentage of the TDA balance collected is significantly greater for self-reported liabilities than when the IRS makes additional assessments. However, AUR assessments result in a greater percentage of dollars collected through offsets.
6. Penalty and interest significantly increase the balance owed by taxpayers, particularly when the underlying balance remains unresolved for several years.
7. The IRS abates between a quarter and a third of TDA liabilities and 40 to 50 percent of its substitute for return assessments. It also abates a high proportion of AUR assessments.
8. The IRS completely resolves most of its TDA modules within the 10-year collection statute, with a resolution rate of about 80 percent for TDAs assigned in 2003 and 2005. Unfortunately, the percent of TDAs resolved has generally declined thereafter. Additionally, the balance owed on these delinquencies has been reduced by less than 50 percent.
9. ACS realizes the largest percentage of TDA balances collected by subsequent payment and offset. While the percentage of dollars abated is high in all TDA collection channels, the abatement rates are significantly higher in the queue and Cff than in ACS. However, even controlling for abatements, ACS collects a greater percentage of the liabilities assigned to it compared to the other TDA functions.³⁵

Possible Future Analyses

We hope to perform a similar analysis on Business Master File (BMF) TDAs. A proper examination of the TDA process must include both IMF and BMF delinquencies. We also want to explore dollars collected and abated, which are generated by IRS additional assessments prior to TDA assignment. Finally, we would like to explore the effect of not only penalty and interest assessments, but also their accruals. In the case of unpaid liabilities, accrued penalties and interest are often never assessed. IRS TDA collections occur within a complex and dynamic environment, and this subject will undoubtedly benefit from many other avenues of study.

³⁵ No active collection occurs on cases in the collection queue; however, offsets still occur and previous IRS notices may continue to generate payments even while the TDA is assigned to the collection queue.

TABLE A-1. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements

Calendar Year	Balance at TDA Assignment	Count of Initial TDAs	Elapsed Years	TDAs Open at End of Year	Initial Balance of TDAs Open at Beginning of Year	Balance of TDAs Open at End of Year	Subsequent Payments	Offsets	Penalty	Interest	Abated
2003	\$15,326,191,192	2,025,755	1	1,558,504	\$12,202,854,742	\$12,321,333,646	(\$1,786,425,808)	(\$611,820,128)	\$89,119,656	\$82,701,617	(\$1,106,542,945)
			2	1,249,051	\$10,705,878,192	\$10,370,295,098	(\$2,953,192,068)	(\$1,004,364,248)	\$269,670,662	\$201,012,229	(\$1,829,914,155)
			3	1,016,996	\$9,603,254,459	\$8,841,346,757	(\$3,801,679,211)	(\$1,277,319,134)	\$585,337,406	\$326,153,660	(\$2,224,684,772)
			4	847,927	\$8,947,277,180	\$7,687,319,274	(\$4,416,806,933)	(\$1,483,618,443)	\$739,002,105	\$452,949,424	(\$2,488,570,542)
			5	714,507	\$8,477,650,276	\$6,877,286,149	(\$4,819,728,794)	(\$1,684,130,906)	\$831,819,282	\$576,698,445	(\$2,684,022,516)
			6	619,011	\$8,148,690,214	\$6,297,481,488	(\$5,073,940,295)	(\$1,829,221,428)	\$889,248,536	\$671,604,646	(\$2,778,370,133)
			7	548,004	\$7,835,654,529	\$5,800,613,947	(\$5,270,506,889)	(\$1,933,592,477)	\$927,727,194	\$753,409,760	(\$2,862,189,720)
			8	483,727	\$7,522,186,532	\$5,380,880,143	(\$5,436,525,749)	(\$2,024,868,988)	\$958,856,995	\$824,597,632	(\$2,919,730,570)
			9	414,757	\$7,139,350,498	\$4,956,881,909	(\$5,577,901,284)	(\$2,101,141,358)	\$976,327,065	\$886,788,555	(\$2,957,500,657)
			10	103,959	\$2,715,523,009	\$1,849,467,741	(\$5,701,245,981)	(\$2,150,746,245)	\$986,076,244	\$1,035,691,869	(\$2,985,977,270)
2004	\$17,476,051,716	2,006,000	1	1,506,593	\$14,153,844,859	\$14,143,137,643	(\$2,501,899,078)	(\$747,349,755)	\$149,376,200	\$105,452,460	(\$1,584,920,917)
			2	1,203,022	\$12,549,660,891	\$11,975,498,939	(\$3,754,805,143)	(\$1,154,990,432)	\$452,808,938	\$226,818,884	(\$2,569,660,336)
			3	992,048	\$11,521,880,241	\$10,436,948,936	(\$4,596,320,388)	(\$1,439,730,392)	\$669,618,362	\$360,478,204	(\$3,187,925,790)
			4	828,892	\$10,878,654,148	\$9,336,449,727	(\$5,148,678,296)	(\$1,702,363,549)	\$801,491,959	\$496,624,078	(\$3,538,338,134)
			5	715,665	\$10,351,952,753	\$8,504,213,714	(\$5,503,793,070)	(\$1,894,349,784)	\$883,250,592	\$608,266,879	(\$3,739,317,881)
			6	633,569	\$10,003,813,674	\$7,901,689,312	(\$5,772,219,032)	(\$2,022,786,186)	\$938,660,871	\$703,039,219	(\$3,862,272,262)
			7	566,207	\$9,700,168,602	\$7,385,215,277	(\$5,988,599,885)	(\$2,135,679,620)	\$982,062,268	\$786,224,962	(\$3,957,635,608)
			8	505,348	\$9,403,151,087	\$6,952,037,321	(\$6,172,225,808)	(\$2,228,208,053)	\$1,007,015,975	\$861,745,767	(\$4,037,658,221)
			9	439,849	\$8,950,989,408	\$6,427,925,814	(\$6,340,470,292)	(\$2,302,166,002)	\$1,037,347,823	\$1,281,890,856	(\$4,094,826,942)
			10	112,388	\$2,958,725,002	\$2,161,305,313	(\$6,460,997,698)	(\$2,354,540,305)	\$1,047,975,390	\$2,042,654,002	(\$4,150,004,992)
2005	\$25,996,084,845	2,607,542	1	1,955,638	\$20,955,172,357	\$20,872,616,207	(\$2,990,840,093)	(\$936,838,872)	\$307,557,342	\$145,216,764	(\$3,513,228,092)
			2	1,592,045	\$18,585,020,738	\$17,657,416,604	(\$4,334,974,344)	(\$1,446,837,231)	\$670,504,187	\$302,527,494	(\$5,507,789,892)
			3	1,326,084	\$17,390,016,757	\$15,759,106,685	(\$5,167,617,507)	(\$1,908,989,084)	\$902,564,484	\$470,456,001	(\$6,399,531,567)
			4	1,146,071	\$16,596,217,970	\$14,423,464,446	(\$5,703,405,817)	(\$2,239,561,550)	\$1,036,694,485	\$620,049,022	(\$6,880,716,807)
			5	1,021,777	\$15,982,934,958	\$13,440,585,575	(\$6,098,077,184)	(\$2,462,184,885)	\$1,125,512,272	\$747,675,302	(\$7,235,932,955)
			6	920,443	\$15,505,712,748	\$12,629,990,619	(\$6,439,393,580)	(\$2,653,438,207)	\$1,195,301,649	\$866,792,903	(\$7,585,208,018)
			7	835,687	\$15,067,568,527	\$11,945,667,325	(\$6,728,867,200)	(\$2,814,797,286)	\$1,238,566,723	\$975,711,954	(\$7,762,861,068)
			8	759,818	\$14,613,428,103	\$11,290,987,893	(\$6,981,187,587)	(\$2,940,382,649)	\$1,267,269,374	\$1,574,319,150	(\$7,934,958,417)
			9	672,606	\$14,138,685,088	\$10,636,264,945	(\$7,190,638,676)	(\$3,041,430,818)	\$1,287,987,437	\$3,395,828,366	(\$8,019,244,730)
			10	509,936	\$10,346,307,042	\$7,790,616,336	(\$7,314,258,218)	(\$3,086,529,497)	\$1,297,081,117	\$3,728,401,948	(\$8,066,761,341)

Table A-1. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements—Continued

Calendar Year	Balance at TDA Assignment	Count of Initial TDAs	Elapsed Years	TDAs Open at End of Year	Initial Balance of TDAs Open at Beginning of Year	Balance of TDAs Open at End of Year	Subsequent Payments	Offsets	Penalty	Interest	Abated
2006	\$30,351,736,552	2,957,590	1	2,203,545	\$24,284,699,486	\$24,382,428,297	(\$3,781,436,023)	(\$1,205,226,894)	\$309,230,189	\$209,887,340	(\$4,028,543,955)
			2	1,780,405	\$22,047,122,624	\$21,134,976,827	(\$5,087,188,574)	(\$1,963,812,846)	\$709,216,745	\$407,185,852	(\$6,082,017,996)
			3	1,509,512	\$20,820,896,493	\$19,211,320,024	(\$5,869,908,534)	(\$2,487,104,515)	\$946,574,128	\$582,951,565	(\$6,954,956,357)
			4	1,328,595	\$19,861,906,645	\$17,796,978,467	(\$6,419,325,485)	(\$2,827,975,036)	\$1,094,243,027	\$732,786,143	(\$7,563,744,203)
			5	1,186,108	\$19,080,518,237	\$16,650,119,657	(\$6,894,884,536)	(\$3,116,809,435)	\$1,212,432,154	\$876,720,366	(\$8,037,931,474)
			6	1,069,254	\$18,450,044,275	\$15,650,453,262	(\$7,275,730,544)	(\$3,348,658,925)	\$1,284,228,928	\$1,003,755,551	(\$8,323,694,663)
			7	972,478	\$17,805,347,394	\$14,741,986,913	(\$7,609,546,792)	(\$3,522,136,383)	\$1,332,557,514	\$1,639,089,778	(\$8,634,838,759)
			8	884,575	\$17,271,468,616	\$13,932,332,131	(\$7,883,737,061)	(\$3,667,649,412)	\$1,364,801,580	\$3,675,241,137	(\$8,832,844,829)
			9	846,817	\$17,053,947,088	\$13,589,243,327	(\$8,056,543,080)	(\$3,741,238,520)	\$1,378,720,590	\$4,248,993,113	(\$8,941,304,377)
			10	846,817	\$17,053,947,088	\$13,589,243,327	(\$8,063,991,533)	(\$3,746,160,529)	\$1,379,211,905	\$4,258,974,459	(\$8,944,717,308)
2007	\$40,678,451,308	3,740,790	1	2,740,824	\$32,849,020,598	\$32,783,285,129	(\$3,664,816,262)	(\$1,704,605,008)	\$409,689,710	\$244,845,100	(\$6,153,964,029)
			2	2,263,145	\$29,935,148,423	\$28,948,279,752	(\$4,995,256,277)	(\$2,578,789,150)	\$928,407,071	\$442,246,287	(\$8,835,026,272)
			3	1,964,827	\$28,301,134,533	\$26,531,718,898	(\$5,902,253,282)	(\$3,123,531,076)	\$1,225,542,214	\$619,234,800	(\$10,248,579,985)
			4	1,740,283	\$26,943,462,391	\$24,561,583,882	(\$6,622,550,732)	(\$3,557,488,509)	\$1,439,733,283	\$784,185,932	(\$11,226,517,587)
			5	1,563,477	\$25,668,240,755	\$22,940,259,164	(\$7,222,864,085)	(\$3,909,144,483)	\$1,564,479,638	\$939,931,827	(\$11,990,004,660)
			6	1,423,312	\$24,806,106,219	\$21,617,558,053	(\$7,740,223,075)	(\$4,171,160,348)	\$1,658,992,874	\$1,578,574,745	(\$12,433,762,035)
			7	1,309,349	\$24,032,849,448	\$20,502,316,740	(\$8,157,636,602)	(\$4,376,959,658)	\$1,722,429,524	\$4,085,368,501	(\$12,878,254,066)
			8	1,266,294	\$23,740,426,822	\$20,017,141,706	(\$8,430,113,249)	(\$4,487,871,926)	\$1,753,835,888	\$4,860,034,842	(\$13,076,870,319)
			9	1,266,294	\$23,740,426,822	\$20,017,141,706	(\$8,437,945,803)	(\$4,493,500,687)	\$1,760,082,415	\$4,884,033,478	(\$13,086,103,480)
			10	1,266,294	\$23,740,426,822	\$20,017,141,706	(\$8,437,945,803)	(\$4,493,500,687)	\$1,760,082,415	\$4,884,033,478	(\$13,086,103,480)
2008	\$36,483,193,583	3,590,831	1	2,649,311	\$29,335,325,217	\$29,514,274,945	(\$3,413,545,809)	(\$1,878,405,331)	\$403,291,535	\$202,209,490	(\$4,823,470,102)
			2	2,208,374	\$26,773,821,008	\$26,230,469,576	(\$4,811,813,396)	(\$2,701,802,914)	\$890,503,910	\$370,626,083	(\$6,864,923,888)
			3	1,901,082	\$24,935,784,088	\$23,876,251,293	(\$5,816,581,381)	(\$3,295,841,646)	\$1,269,411,583	\$528,307,655	(\$8,035,065,577)
			4	1,670,837	\$23,366,011,300	\$21,993,850,951	(\$6,600,450,645)	(\$3,744,521,767)	\$1,495,474,859	\$681,365,525	(\$8,780,833,412)
			5	1,496,220	\$22,207,414,517	\$20,447,570,975	(\$7,255,716,053)	(\$4,064,995,130)	\$1,632,099,721	\$1,173,686,878	(\$9,281,979,287)
			6	1,359,759	\$21,073,271,206	\$18,980,116,872	(\$7,761,069,616)	(\$4,306,986,432)	\$1,720,257,087	\$2,873,881,178	(\$9,741,262,882)
			7	1,312,693	\$20,740,005,590	\$18,478,220,702	(\$8,046,513,063)	(\$4,411,125,862)	\$1,755,943,129	\$3,458,126,788	(\$9,937,104,100)
			8	1,312,693	\$20,740,005,590	\$18,478,220,702	(\$8,042,788,704)	(\$4,413,244,389)	\$1,756,334,243	\$3,461,084,068	(\$9,939,486,398)
			9	1,312,693	\$20,740,005,590	\$18,478,220,702	(\$8,042,788,704)	(\$4,413,244,389)	\$1,756,334,243	\$3,461,084,068	(\$9,939,486,398)
			10	1,312,693	\$20,740,005,590	\$18,478,220,702	(\$8,042,788,704)	(\$4,413,244,389)	\$1,756,334,243	\$3,461,084,068	(\$9,939,486,398)

Table A-1. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements—Continued

Calendar Year	Balance at TDA Assignment	Count of Initial TDAs	Elapsed Years	TDAs Open at End of Year	Initial Balance of TDAs Open at Beginning of Year	Balance of TDAs Open at End of Year	Subsequent Payments	Offsets	Penalty	Interest	Abated
2009	\$41,987,700,518	3,625,853	1	2,855,031	\$34,910,079,884	\$35,332,522,291	(\$3,631,938,674)	(\$1,725,467,681)	\$534,920,665	\$153,433,427	(\$5,003,814,841)
			2	2,436,301	\$31,718,637,329	\$31,581,177,384	(\$5,307,444,529)	(\$2,621,430,430)	\$1,107,107,550	\$296,769,692	(\$7,490,662,803)
			3	2,132,368	\$29,367,116,987	\$28,767,291,989	(\$6,524,263,263)	(\$3,251,623,753)	\$1,561,792,804	\$447,222,002	(\$8,852,151,280)
			4	1,902,942	\$27,478,019,214	\$26,495,824,596	(\$7,469,086,900)	(\$3,698,507,804)	\$1,824,680,818	\$951,475,433	(\$9,784,258,243)
			5	1,727,894	\$26,092,408,873	\$24,608,525,544	(\$8,215,663,586)	(\$4,037,187,165)	\$1,996,145,246	\$2,667,410,510	(\$10,433,401,857)
			6	1,674,774	\$25,649,101,721	\$23,922,424,354	(\$8,595,196,469)	(\$4,171,691,809)	\$2,071,626,158	\$3,234,730,639	(\$10,715,471,960)
			7	1,674,774	\$25,649,101,721	\$23,922,424,354	(\$8,598,477,760)	(\$4,173,553,636)	\$2,071,942,891	\$3,236,409,717	(\$10,716,623,485)
			8	1,674,774	\$25,649,101,721	\$23,922,424,354	(\$8,598,477,760)	(\$4,173,553,636)	\$2,071,942,891	\$3,236,409,717	(\$10,716,623,485)
			9	1,674,774	\$25,649,101,721	\$23,922,424,354	(\$8,598,477,760)	(\$4,173,553,636)	\$2,071,942,891	\$3,236,409,717	(\$10,716,623,485)
			10	1,674,774	\$25,649,101,721	\$23,922,424,354	(\$8,598,477,760)	(\$4,173,553,636)	\$2,071,942,891	\$3,236,409,717	(\$10,716,623,485)
2010	\$45,704,392,188	4,150,283	1	3,104,902	\$36,225,946,333	\$36,770,145,599	(\$3,900,220,431)	(\$2,003,251,896)	\$577,836,901	\$145,097,158	(\$6,856,757,918)
			2	2,576,011	\$32,624,372,420	\$32,514,634,453	(\$5,661,580,107)	(\$2,941,601,479)	\$1,178,134,280	\$281,162,628	(\$9,488,141,786)
			3	2,231,853	\$29,863,854,205	\$29,468,592,163	(\$6,925,540,407)	(\$3,556,525,164)	\$1,676,116,937	\$652,063,333	(\$11,204,940,751)
			4	1,984,379	\$28,100,821,406	\$27,193,787,060	(\$7,871,884,638)	(\$4,003,121,592)	\$1,948,303,588	\$1,942,072,455	(\$12,079,326,237)
			5	1,907,417	\$27,527,977,964	\$26,369,110,984	(\$8,443,668,673)	(\$4,228,019,965)	\$2,071,093,184	\$2,535,673,602	(\$12,484,578,501)
			6	1,907,417	\$27,527,977,964	\$26,369,110,984	(\$8,460,865,513)	(\$4,240,011,358)	\$2,073,829,920	\$2,545,381,557	(\$12,492,491,633)
			7	1,907,417	\$27,527,977,964	\$26,369,110,984	(\$8,460,865,513)	(\$4,240,011,358)	\$2,073,829,920	\$2,545,381,557	(\$12,492,491,633)
			8	1,907,417	\$27,527,977,964	\$26,369,110,984	(\$8,460,865,513)	(\$4,240,011,358)	\$2,073,829,920	\$2,545,381,557	(\$12,492,491,633)
			9	1,907,417	\$27,527,977,964	\$26,369,110,984	(\$8,460,865,513)	(\$4,240,011,358)	\$2,073,829,920	\$2,545,381,557	(\$12,492,491,633)
			10	1,907,417	\$27,527,977,964	\$26,369,110,984	(\$8,460,865,513)	(\$4,240,011,358)	\$2,073,829,920	\$2,545,381,557	(\$12,492,491,633)
2011	\$42,926,217,917	4,024,360	1	3,020,150	\$34,032,316,693	\$34,795,749,849	(\$3,800,144,148)	(\$1,797,902,891)	\$516,330,747	\$126,728,719	(\$6,755,704,392)
			2	2,522,271	\$29,319,068,318	\$29,792,646,489	(\$5,548,247,612)	(\$2,676,941,031)	\$1,078,438,736	\$387,516,960	(\$10,289,591,480)
			3	2,196,229	\$27,055,074,385	\$27,132,375,863	(\$6,725,846,404)	(\$3,273,238,113)	\$1,511,675,025	\$1,265,063,880	(\$11,484,501,895)
			4	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,414,305,809)	(\$3,569,753,914)	\$1,691,974,661	\$1,749,969,482	(\$11,981,924,704)
			5	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,435,121,689)	(\$3,583,172,056)	\$1,697,342,413	\$1,760,693,864	(\$11,990,870,525)
			6	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,435,121,689)	(\$3,583,172,056)	\$1,697,342,413	\$1,760,693,864	(\$11,990,870,525)
			7	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,435,121,689)	(\$3,583,172,056)	\$1,697,342,413	\$1,760,693,864	(\$11,990,870,525)
			8	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,435,121,689)	(\$3,583,172,056)	\$1,697,342,413	\$1,760,693,864	(\$11,990,870,525)
			9	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,435,121,689)	(\$3,583,172,056)	\$1,697,342,413	\$1,760,693,864	(\$11,990,870,525)
			10	2,097,838	\$26,304,398,711	\$26,094,655,803	(\$7,435,121,689)	(\$3,583,172,056)	\$1,697,342,413	\$1,760,693,864	(\$11,990,870,525)

Table A-1. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements—Continued

Calendar Year	Balance at TDA Assignment	Count of Initial TDAs	Elapsed Years	TDAs Open at End of Year	Initial Balance of TDAs Open at Beginning of Year	Balance of TDAs Open at End of Year	Subsequent Payments	Offsets	Penalty	Interest	Abated
2012	\$40,571,783.915	3,958,801	1	3,042,337	\$32,742,725.029	\$33,739,786.846	(\$3,603,072,713)	(\$1,763,779,947)	\$543,964,069	\$180,683,705	(\$5,183,970,519)
			2	2,560,958	\$28,855,402,233	\$29,385,315,074	(\$5,293,758,985)	(\$2,675,662,462)	\$1,075,108,307	\$758,149,918	(\$7,945,451,495)
			3	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,264,144,364)	(\$3,108,350,689)	\$1,401,487,635	\$1,269,916,084	(\$8,842,682,967)
			4	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)
			5	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)
			6	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)
			7	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)
			8	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)
			9	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)
			10	2,431,227	\$27,700,178,097	\$28,084,991,830	(\$6,290,918,513)	(\$3,122,370,872)	\$1,405,404,179	\$1,276,993,550	(\$8,856,749,351)

TABLE A-2. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Balance Due Ranges

Calendar Year	Description	\$1 to \$1,000	\$1,001 to \$2,000	\$2,001 to \$5,000	\$5,001 to \$10,000	\$10,001 to \$25,000	Greater Than \$25,000
2003	TDA Count	451,712	505,146	565,164	250,331	160,431	92,971
	Initial Balance Due	\$240,828,346	\$740,649,611	\$1,793,135,994	\$1,745,405,984	\$2,446,859,748	\$8,359,311,510
	Subsequent Payments	(\$159,645,920)	(\$359,532,161)	(\$886,819,407)	(\$845,113,111)	(\$1,054,887,316)	(\$2,395,248,066)
	Offset	(\$120,551,135)	(\$323,517,417)	(\$625,523,335)	(\$382,757,722)	(\$316,904,898)	(\$381,491,738)
	Penalty	\$33,446,962	\$78,987,494	\$172,236,063	\$141,164,346	\$168,647,485	\$391,593,894
	Interest	\$45,822,418	\$91,605,397	\$204,538,984	\$171,753,801	\$186,241,347	\$335,729,923
	Abated	(\$27,475,487)	(\$43,593,443)	(\$146,029,262)	(\$188,258,429)	(\$333,475,486)	(\$2,247,145,162)
2004	TDA Count	423,322	454,156	576,996	259,361	175,488	116,677
	Initial Balance Due	\$226,310,020	\$677,561,180	\$1,825,708,766	\$1,813,925,576	\$2,685,068,369	\$10,247,477,805
	Subsequent Payments	(\$175,904,106)	(\$375,205,472)	(\$982,626,633)	(\$934,209,391)	(\$1,225,047,057)	(\$2,768,005,039)
	Offset	(\$121,702,024)	(\$321,705,919)	(\$694,321,067)	(\$425,213,207)	(\$369,848,986)	(\$421,749,103)
	Penalty	\$31,504,743	\$75,081,251	\$176,833,442	\$144,534,463	\$180,793,631	\$439,227,860
	Interest	\$49,458,179	\$105,365,867	\$277,857,329	\$269,972,536	\$358,344,029	\$981,656,061
	Abated	(\$23,945,856)	(\$47,587,534)	(\$162,623,959)	(\$224,236,892)	(\$456,718,448)	(\$3,234,892,303)
2005	TDA Count	467,988	561,662	762,610	388,628	254,399	172,255
	Initial Balance Due	\$250,772,306	\$832,921,777	\$2,462,466,838	\$2,713,904,406	\$3,886,596,037	\$15,849,423,481
	Subsequent Payments	(\$197,408,638)	(\$453,485,724)	(\$1,226,197,456)	(\$1,188,786,059)	(\$1,422,684,108)	(\$2,825,696,234)
	Offset	(\$138,643,683)	(\$412,526,218)	(\$935,264,928)	(\$616,224,119)	(\$496,862,712)	(\$487,007,838)
	Penalty	\$35,438,062	\$86,715,864	\$216,494,334	\$189,194,145	\$232,769,822	\$536,468,890
	Interest	\$59,999,279	\$140,495,631	\$436,422,558	\$502,988,977	\$683,406,500	\$1,905,089,002
	Abated	(\$30,966,042)	(\$62,829,089)	(\$266,610,631)	(\$409,931,148)	(\$766,502,708)	(\$6,529,921,723)
2006	TDA Count	509,337	615,280	867,067	474,879	299,498	191,529
	Initial Balance Due	\$271,505,543	\$909,001,299	\$2,834,634,772	\$3,307,189,504	\$4,562,501,833	\$18,466,903,602
	Subsequent Payments	(\$252,565,017)	(\$472,447,644)	(\$1,296,013,518)	(\$1,308,078,558)	(\$1,546,630,399)	(\$3,188,256,398)
	Offset	(\$150,453,775)	(\$458,496,094)	(\$1,109,332,469)	(\$813,619,177)	(\$606,206,985)	(\$608,052,027)
	Penalty	\$36,272,026	\$87,550,652	\$235,367,768	\$217,417,225	\$250,427,660	\$552,176,574
	Interest	\$69,173,136	\$151,472,115	\$476,431,996	\$576,607,004	\$768,858,956	\$2,216,431,253
	Abated	(\$46,693,380)	(\$77,413,359)	(\$308,977,379)	(\$480,897,741)	(\$847,379,103)	(\$7,183,356,346)

TABLE A-2. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Balance Due Ranges—Continued

Calendar Year	Description	\$1 to \$1,000	\$1,001 to \$2,000	\$2,001 to \$5,000	\$5,001 to \$10,000	\$10,001 to \$25,000	Greater Than \$25,000
2007	TDA Count	781,534	666,064	1,006,717	616,892	408,744	260,839
	Initial Balance Due	\$449,269,937	\$978,486,020	\$3,313,143,007	\$4,309,521,457	\$6,214,721,121	\$25,413,309,766
	Subsequent Payments	(\$271,255,779)	(\$440,895,663)	(\$1,317,912,732)	(\$1,436,890,739)	(\$1,656,381,839)	(\$3,314,609,051)
	Offset	(\$275,425,703)	(\$501,704,901)	(\$1,240,557,318)	(\$1,002,620,280)	(\$735,452,110)	(\$737,740,375)
	Penalty	\$52,681,418	\$88,781,235	\$249,790,312	\$252,451,642	\$304,732,788	\$811,645,020
	Interest	\$86,193,479	\$145,245,793	\$486,817,479	\$630,865,633	\$884,120,071	\$2,650,791,023
	Abated	(\$129,967,848)	(\$94,765,806)	(\$402,625,097)	(\$679,848,856)	(\$1,203,496,891)	(\$10,575,398,982)
2008	TDA Count	731,451	670,937	1,016,335	550,220	385,289	236,599
	Initial Balance Due	\$414,830,731	\$990,497,117	\$3,284,041,388	\$3,845,368,753	\$5,906,630,325	\$22,041,825,269
	Subsequent Payments	(\$233,650,205)	(\$403,601,852)	(\$1,205,681,275)	(\$1,232,376,624)	(\$1,579,199,205)	(\$3,388,279,541)
	Offset	(\$243,479,900)	(\$497,742,738)	(\$1,214,494,488)	(\$881,503,108)	(\$721,414,352)	(\$854,609,803)
	Penalty	\$48,985,319	\$87,949,930	\$244,002,549	\$246,084,994	\$317,910,922	\$811,400,529
	Interest	\$69,958,224	\$121,666,892	\$380,821,867	\$454,035,901	\$670,573,651	\$1,764,027,532
	Abated	(\$74,950,472)	(\$84,664,150)	(\$337,726,625)	(\$498,468,833)	(\$991,149,552)	(\$7,952,526,766)
2009	TDA Count	520,936	596,584	1,038,156	697,680	479,893	292,604
	Initial Balance Due	\$290,826,653	\$907,622,403	\$3,388,475,005	\$4,874,905,595	\$7,345,952,997	\$25,179,917,864
	Subsequent Payments	(\$168,308,125)	(\$365,851,165)	(\$1,169,531,683)	(\$1,324,853,213)	(\$1,714,583,628)	(\$3,855,349,945)
	Offset	(\$133,042,345)	(\$362,615,410)	(\$1,039,704,190)	(\$936,893,572)	(\$753,063,838)	(\$948,234,281)
	Penalty	\$35,240,420	\$77,982,626	\$246,319,981	\$277,035,867	\$375,311,428	\$1,060,052,569
	Interest	\$43,615,381	\$87,002,543	\$304,084,837	\$412,359,663	\$623,590,625	\$1,765,756,667
	Abated	(\$58,131,912)	(\$79,760,319)	(\$365,636,375)	(\$663,016,312)	(\$1,322,675,570)	(\$8,227,402,996)
2010	TDA Count	840,148	808,468	1,121,844	647,137	453,016	279,670
	Initial Balance Due	\$490,472,989	\$1,199,142,789	\$3,618,771,867	\$4,555,543,493	\$6,918,572,956	\$28,921,888,093
	Subsequent Payments	(\$254,834,586)	(\$456,192,529)	(\$1,168,128,274)	(\$1,194,754,484)	(\$1,618,245,117)	(\$3,768,710,523)
	Offset	(\$240,509,528)	(\$508,265,653)	(\$1,143,358,182)	(\$810,941,332)	(\$678,453,146)	(\$858,483,517)
	Penalty	\$56,440,339	\$101,938,493	\$258,977,602	\$252,636,859	\$349,515,692	\$1,054,320,934
	Interest	\$47,796,130	\$86,595,116	\$244,066,722	\$302,628,434	\$445,969,685	\$1,418,325,469
	Abated	(\$51,894,751)	(\$87,114,213)	(\$305,230,442)	(\$536,651,544)	(\$1,096,370,320)	(\$10,415,230,363)

TABLE A-2. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Balance Due Ranges—Continued

Calendar Year	Description	\$1 to \$1,000	\$1,001 to \$2,000	\$2,001 to \$5,000	\$5,001 to \$10,000	\$10,001 to \$25,000	Greater Than \$25,000
2011	TDA Count	825,154	754,679	1,136,688	639,600	422,102	246,137
	Initial Balance Due	\$480,421,472	\$1,117,857,159	\$3,718,666,067	\$4,484,256,769	\$6,436,680,632	\$26,688,335,819
	Subsequent Payments	(\$218,477,840)	(\$383,128,439)	(\$1,047,791,886)	(\$1,060,125,297)	(\$1,423,029,355)	(\$3,302,568,871)
	Offset	(\$224,505,778)	(\$430,557,063)	(\$1,011,715,750)	(\$706,990,836)	(\$567,489,241)	(\$641,913,390)
	Penalty	\$46,648,243	\$81,026,407	\$226,937,654	\$209,268,186	\$285,370,582	\$848,091,340
	Interest	\$36,845,747	\$62,819,645	\$182,749,854	\$215,524,482	\$308,187,021	\$954,567,115
	Abated	(\$78,541,023)	(\$83,847,322)	(\$322,574,937)	(\$595,578,698)	(\$922,707,731)	(\$9,987,620,814)
	TDA Count	797,290	747,214	1,171,478	612,676	396,654	233,489
2012	Initial Balance Due	\$461,112,482	\$1,110,719,485	\$3,825,438,842	\$4,282,481,148	\$5,990,784,717	\$24,901,247,241
	Subsequent Payments	(\$177,373,252)	(\$303,891,260)	(\$820,145,806)	(\$813,093,192)	(\$1,122,292,940)	(\$3,054,122,063)
	Offset	(\$179,411,721)	(\$380,643,270)	(\$941,082,680)	(\$607,840,547)	(\$465,246,196)	(\$548,146,459)
	Penalty	\$39,838,543	\$63,657,332	\$175,282,006	\$165,586,345	\$227,507,301	\$733,532,652
	Interest	\$25,497,597	\$42,359,620	\$121,513,641	\$136,305,066	\$198,745,472	\$752,572,154
	Abated	(\$57,897,841)	(\$69,743,692)	(\$281,809,666)	(\$399,235,354)	(\$749,409,440)	(\$7,298,653,359)

TABLE A-3. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Source of Assessment

Calendar Year	Description	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2003	TDA Count	76,305	127,153	248,863	1,115,474	110,324
	Initial Balance Due	\$2,755,702,547	\$1,641,819,169	\$955,455,897	\$5,935,437,250	\$1,636,660,799
	Subsequent Payments	(\$397,863,606)	(\$381,301,366)	(\$317,247,600)	(\$3,352,551,156)	(\$267,458,821)
	Offset	(\$109,821,313)	(\$198,452,772)	(\$323,382,148)	(\$1,076,177,894)	(\$103,027,521)
	Penalty	\$57,352,147	\$104,629,310	\$87,347,519	\$594,619,903	\$34,505
	Interest	\$66,142,201	\$72,386,963	\$65,165,077	\$542,889,882	\$115,653,487
	Abated	(\$1,347,079,999)	(\$245,313,974)	(\$145,199,035)	(\$362,429,282)	(\$638,341,982)
	TDA Count	148,322	130,994	256,753	1,028,470	114,071
2004	Initial Balance Due	\$5,064,578,631	\$1,790,767,383	\$1,096,964,103	\$5,585,456,769	\$1,767,559,350
	Subsequent Payments	(\$784,223,463)	(\$474,730,184)	(\$366,381,883)	(\$3,505,764,852)	(\$289,675,909)
	Offset	(\$229,042,168)	(\$220,626,090)	(\$395,625,962)	(\$1,055,677,394)	(\$106,505,103)
	Penalty	\$126,340,535	\$120,222,356	\$95,416,239	\$556,173,822	\$29,803
	Interest	\$453,442,683	\$182,694,977	\$108,691,313	\$728,528,130	\$296,995,659
	Abated	(\$2,340,659,814)	(\$231,996,225)	(\$232,552,410)	(\$380,021,768)	(\$767,861,296)
	TDA Count	512,705	224,899	301,903	1,104,554	101,588
	Initial Balance Due	\$13,362,507,090	\$2,034,842,462	\$1,347,565,321	\$5,544,297,832	\$1,500,995,601
2005	Subsequent Payments	(\$1,719,278,819)	(\$563,315,839)	(\$423,463,099)	(\$3,328,886,959)	(\$248,869,759)
	Offset	(\$682,672,073)	(\$398,246,119)	(\$431,513,171)	(\$1,096,714,008)	(\$94,913,229)
	Penalty	\$372,356,726	\$150,049,224	\$102,910,162	\$534,072,724	\$45,906
	Interest	\$1,754,673,244	\$272,574,120	\$140,684,401	\$901,732,852	\$294,428,237
	Abated	(\$6,251,321,458)	(\$240,489,498)	(\$384,254,445)	(\$338,857,989)	(\$604,296,121)

TABLE A-3. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Source of Assessment—Continued

Calendar Year	Description	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2006	TDA Count	588,943	273,576	374,373	1,214,361	102,374
	Initial Balance Due	\$15,272,022,052.9	\$2,699,929,404.0	\$1,546,702,949.2	\$6,662,831,209.9	\$1,555,671,271.1
	Subsequent Payments	(\$1,657,436,219.2)	(\$708,090,013.6)	(\$483,120,497.2)	(\$3,904,703,877.4)	(\$220,079,883.5)
	Offset	(\$798,615,902.5)	(\$568,757,077.8)	(\$559,941,489.1)	(\$1,279,736,554.8)	(\$96,242,450.3)
	Penalty	\$313,100,522.7	\$185,710,811.0	\$110,596,640.4	\$616,339,300.9	\$35,150.0
	Interest	\$2,000,147,374.8	\$396,544,510.4	\$161,928,187.0	\$1,022,051,805.4	\$282,864,102.1
	Abated	(\$6,571,418,753.8)	(\$368,654,726.7)	(\$396,185,415.2)	(\$576,039,465.8)	(\$641,631,939.2)
2007	TDA Count	1,043,966	328,809	491,988	1,330,059	104,034
	Initial Balance Due	\$24,497,035,199.50	\$2,852,164,654.01	\$1,874,553,318.62	\$7,150,372,177.98	\$1,471,575,406.52
	Subsequent Payments	(\$2,406,444,052.86)	(\$674,618,290.74)	(\$476,662,500.63)	(\$3,656,970,935.55)	(\$173,447,577.30)
	Offset	(\$1,135,157,789.35)	(\$718,330,580.74)	(\$679,509,112.51)	(\$1,396,911,300.99)	(\$92,287,027.64)
	Penalty	\$559,551,263.81	\$219,753,090.56	\$127,434,963.11	\$679,624,580.41	\$36,854.73
	Interest	\$2,693,063,064.29	\$396,338,868.59	\$163,748,874.45	\$1,008,328,975.43	\$226,404,885.33
	Abated	(\$10,482,352,227.22)	(\$397,787,490.46)	(\$532,432,329.44)	(\$826,537,704.52)	(\$509,468,588.88)
2008	TDA Count	739,496	311,267	544,576	1,408,584	126,674
	Initial Balance Due	\$17,215,158,029	\$3,275,957,777	\$2,464,789,305	\$8,573,658,078	\$1,692,197,368
	Subsequent Payments	(\$1,549,671,652)	(\$746,128,193)	(\$554,527,916)	(\$3,927,091,666)	(\$171,233,155)
	Offset	(\$730,838,818)	(\$665,223,976)	(\$827,850,704)	(\$1,589,699,028)	(\$103,947,144)
	Penalty	\$337,635,916	\$238,267,045	\$159,606,831	\$821,717,247	\$35,763
	Interest	\$1,439,847,831	\$332,816,518	\$163,867,639	\$955,104,268	\$203,199,971
	Abated	(\$6,728,095,595)	(\$525,851,177)	(\$747,290,401)	(\$994,646,932)	(\$524,877,644)

TABLE A-3. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Source of Assessment—Continued

Calendar Year	Description	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2009	TDA Count	1,059,119	328,596	421,222	1,258,574	133,484
	Initial Balance Due	\$21,519,252,921	\$3,510,915,440	\$2,190,381,922	\$9,287,978,110	\$1,791,857,278
	Subsequent Payments	(\$1,937,916,493)	(\$734,096,173)	(\$527,178,242)	(\$4,164,607,080)	(\$166,332,122)
	Offset	(\$891,601,842)	(\$695,800,264)	(\$614,982,034)	(\$1,417,235,663)	(\$107,247,687)
	Penalty	\$573,265,383	\$267,037,832	\$135,425,080	\$878,486,072	\$33,808
	Interest	\$1,515,924,103	\$301,249,911	\$118,011,988	\$803,810,795	\$172,133,849
	Abated	(\$7,729,519,040)	(\$449,702,771)	(\$599,496,724)	(\$822,895,861)	(\$496,312,590)
2010	TDA Count	786,668	341,242	696,468	1,615,298	165,008
	Initial Balance Due	\$21,846,637,985	\$3,745,192,392	\$3,213,707,429	\$10,254,615,336	\$2,202,121,920
	Subsequent Payments	(\$1,264,503,773)	(\$663,033,824)	(\$672,146,619)	(\$4,465,936,541)	(\$191,300,255)
	Offset	(\$577,674,767)	(\$599,513,270)	(\$879,157,918)	(\$1,546,383,735)	(\$123,339,260)
	Penalty	\$450,454,430	\$295,161,887	\$191,439,933	\$896,627,625	\$39,977
	Interest	\$1,041,306,173	\$234,136,935	\$130,949,018	\$694,263,873	\$150,003,036
	Abated	(\$8,523,041,119)	(\$762,864,033)	(\$896,730,683)	(\$1,006,484,824)	(\$684,970,705)
2011	TDA Count	769,554	320,692	703,735	1,473,234	206,428
	Initial Balance Due	\$19,860,173,613	\$3,860,505,398	\$2,467,419,131	\$9,734,400,489	\$2,510,939,801
	Subsequent Payments	(\$1,301,107,710)	(\$560,341,761)	(\$518,971,847)	(\$3,893,835,651)	(\$194,376,899)
	Offset	(\$532,457,391)	(\$545,391,380)	(\$766,601,880)	(\$1,198,087,537)	(\$124,133,401)
	Penalty	\$403,543,321	\$214,594,684	\$138,486,517	\$734,750,156	\$38,980
	Interest	\$659,627,735	\$159,112,330	\$89,564,038	\$509,380,819	\$127,747,972
	Abated	(\$7,927,276,034)	(\$740,644,539)	(\$435,026,997)	(\$1,534,273,270)	(\$731,930,494)

Table A-3. TDA Modules, Balances Due, Dollars Collected by Subsequent Payments and Offsets, Assessed Penalties, Assessed Interest, and Abatements by Source of Assessment—Continued

Calendar Year	Description	Self-Reported Assessments	Substitute for Return	Audit Assessments	AUR Assessments	Trust Fund Recovery Penalties
2012	TDA Count	530,422	346,807	667,103	1,601,048	217,242
	Initial Bal. Due	\$15,981,125,775	\$4,232,164,705	\$2,991,886,015	\$10,224,391,491	\$2,483,987,382
	Sub. Payment	(\$799,485,420)	(\$563,383,176)	(\$443,036,348)	(\$3,554,387,916)	(\$166,863,415)
	Offset	(\$344,985,849)	(\$507,554,114)	(\$752,560,887)	(\$1,052,606,719)	(\$105,149,706)
	Penalty	\$286,208,504	\$212,280,910	\$126,427,212	\$623,128,493	\$39,077
	Interest	\$417,592,869	\$147,703,005	\$66,177,617	\$405,922,275	\$85,004,578
	Abated	(\$5,860,055,257)	(\$341,658,757)	(\$681,903,277)	(\$635,580,165)	(\$679,545,902)

TABLE A-4. Initial TDA Balance, Subsequent Payments, Offsets, and Abatements by Collection Channel*

ACS				
Calendar Year	Initial Balance Due	Subsequent Payments	Offsets	Abated
2003	\$7,792,592,325	(\$3,426,144,186)	(\$1,700,612,873)	(\$1,101,444,823)
2004	\$8,055,134,988	(\$3,751,122,687)	(\$1,821,135,021)	(\$1,374,908,979)
2005	\$10,998,087,606	(\$4,449,976,986)	(\$2,306,307,552)	(\$2,323,868,875)
2006	\$11,745,756,134	(\$4,958,995,889)	(\$2,669,338,955)	(\$2,231,454,323)
2007	\$13,328,119,659	(\$5,152,715,921)	(\$3,313,012,446)	(\$2,498,865,753)
2008	\$13,076,613,620	(\$4,952,000,018)	(\$3,342,342,605)	(\$2,005,516,405)
2009	\$20,164,274,356	(\$6,033,827,223)	(\$3,225,236,763)	(\$4,106,056,899)
2010	\$23,890,067,756	(\$6,504,108,404)	(\$3,601,310,254)	(\$4,345,387,578)
2011	\$20,559,657,101	(\$5,362,106,864)	(\$3,035,428,058)	(\$3,362,113,103)
2012	\$15,766,253,590	(\$3,680,718,002)	(\$2,554,868,769)	(\$1,949,706,639)
Queue				
Calendar Year	Initial Balance Due	Subsequent Payments	Offsets	Abated
2003	\$4,456,531,893	(\$1,302,443,755)	(\$274,152,689)	(\$1,025,099,018)
2004	\$5,251,622,031	(\$1,413,913,762)	(\$303,798,101)	(\$1,356,746,265)
2005	\$7,259,341,395	(\$1,236,407,732)	(\$339,609,532)	(\$2,698,940,436)
2006	\$10,364,534,372	(\$1,705,896,251)	(\$569,806,186)	(\$3,872,105,194)
2007	\$13,356,607,079	(\$1,474,213,610)	(\$514,373,032)	(\$4,802,778,031)
2008	\$11,887,839,882	(\$1,544,260,843)	(\$558,938,774)	(\$3,713,488,466)
2009	\$9,028,536,600	(\$853,994,321)	(\$364,974,713)	(\$2,467,988,439)
2010	\$14,770,625,847	(\$1,165,220,750)	(\$443,402,944)	(\$5,340,619,441)
2011	\$15,017,679,946	(\$1,167,520,082)	(\$382,384,896)	(\$5,536,502,040)
2012	\$16,502,893,644	(\$1,501,522,558)	(\$413,388,933)	(\$4,112,045,801)
Cff				
Calendar Year	Initial Balance Due	Subsequent Payments	Offsets	Abated
2003	\$3,077,066,975	(\$972,658,039)	(\$175,980,683)	(\$859,433,429)
2004	\$4,169,294,696	(\$1,295,961,250)	(\$229,607,183)	(\$1,418,349,748)
2005	\$7,738,655,844	(\$1,627,873,501)	(\$440,612,413)	(\$3,043,952,030)
2006	\$8,241,446,047	(\$1,399,099,392)	(\$507,015,387)	(\$2,841,157,792)
2007	\$13,993,724,570	(\$1,811,016,272)	(\$666,115,209)	(\$5,784,459,696)
2008	\$11,518,740,081	(\$1,546,527,842)	(\$511,963,009)	(\$4,220,481,526)
2009	\$12,794,889,563	(\$1,710,656,216)	(\$583,342,161)	(\$4,142,578,147)
2010	\$7,043,698,585	(\$791,536,359)	(\$195,298,160)	(\$2,806,484,614)
2011	\$7,348,880,870	(\$905,494,742)	(\$165,359,102)	(\$3,092,255,382)
2012	\$8,302,636,681	(\$1,108,677,954)	(\$154,113,170)	(\$2,794,996,912)

*NOTE: The IRS is required by law to write off any remaining balance due at the expiration of the collection statute of limitations period (generally 10 years from the date of liability assessment, but this period may be extended for several reasons, including bankruptcy).

TABLE A-5. Initial TDA Balance, Subsequent Payments, Offsets, and Abatements by Collection Channel After Six Years*

ACS				
Calendar Year	Initial Balance Due	Subsequent Payments	Offsets	Abated
2003	\$7,792,592,325	(\$3,073,090,254)	(\$1,462,600,955)	(\$1,027,367,242)
2004	\$8,055,134,988	(\$3,401,366,362)	(\$1,591,054,713)	(\$1,298,468,755)
2005	\$10,998,087,606	(\$4,033,340,814)	(\$2,034,143,955)	(\$2,201,007,908)
2006	\$11,745,756,134	(\$4,582,308,923)	(\$2,433,892,394)	(\$2,123,881,197)
2007	\$13,328,119,659	(\$4,848,687,675)	(\$3,125,283,670)	(\$2,413,837,112)
2008	\$13,076,613,620	(\$4,814,247,019)	(\$3,277,483,164)	(\$1,979,716,054)
2009	\$20,164,274,356	(\$6,032,295,430)	(\$3,223,772,255)	(\$4,105,339,235)
2010	\$23,890,067,756	(\$6,504,108,404)	(\$3,601,310,254)	(\$4,345,387,578)
2011	\$20,559,657,101	(\$5,362,106,864)	(\$3,035,428,058)	(\$3,362,113,103)
2012	\$15,766,253,590	(\$3,680,718,002)	(\$2,554,868,769)	(\$1,949,706,639)
Queue				
Calendar Year	Initial Balance Due	Subsequent Payments	Offsets	Abated
2003	\$4,456,531,893	(\$1,133,286,932)	(\$221,764,929)	(\$947,606,322)
2004	\$5,251,622,031	(\$1,227,076,427)	(\$246,025,904)	(\$1,261,716,412)
2005	\$7,259,341,395	(\$1,030,055,031)	(\$271,314,994)	(\$2,534,957,190)
2006	\$10,364,534,372	(\$1,505,558,020)	(\$492,056,484)	(\$3,568,299,367)
2007	\$13,356,607,079	(\$1,291,060,172)	(\$456,578,391)	(\$4,566,614,755)
2008	\$11,887,839,882	(\$1,472,606,814)	(\$537,914,978)	(\$3,625,718,411)
2009	\$9,028,536,600	(\$853,546,234)	(\$364,865,502)	(\$2,467,952,050)
2010	\$14,770,625,847	(\$1,165,220,750)	(\$443,402,944)	(\$5,340,619,441)
2011	\$15,017,679,946	(\$1,167,520,082)	(\$382,384,896)	(\$5,536,502,040)
2012	\$16,502,893,644	(\$1,501,522,558)	(\$413,388,933)	(\$4,112,045,801)
Cff				
Calendar Year	Initial Balance Due	Subsequent Payments	Offsets	Abated
2003	\$3,077,066,975	(\$867,563,109)	(\$144,855,544)	(\$803,396,569)
2004	\$4,169,294,696	(\$1,143,776,243)	(\$185,705,568)	(\$1,302,087,094)
2005	\$7,738,655,844	(\$1,375,997,735)	(\$347,979,258)	(\$2,849,242,920)
2006	\$8,241,446,047	(\$1,187,863,602)	(\$422,710,047)	(\$2,631,514,099)
2007	\$13,993,724,570	(\$1,600,475,227)	(\$589,298,287)	(\$5,453,310,168)
2008	\$11,518,740,081	(\$1,474,215,783)	(\$491,588,290)	(\$4,135,828,417)
2009	\$12,794,889,563	(\$1,709,354,805)	(\$583,054,052)	(\$4,142,180,675)
2010	\$7,043,698,585	(\$791,536,359)	(\$195,298,160)	(\$2,806,484,614)
2011	\$7,348,880,870	(\$905,494,742)	(\$165,359,102)	(\$3,092,255,382)
2012	\$8,302,636,681	(\$1,108,677,954)	(\$154,113,170)	(\$2,794,996,912)

*NOTE: The IRS is required by law to write off any remaining balance due at the expiration of the collection statute of limitations period (generally 10 years from the date of liability assessment, but this period may be extended for several reasons, including bankruptcy).

Analysis of Flow-Through Entities Using Social Network Analysis Techniques

*Ashish Agarwal and Shannon Chen, McCombs School of Business, University of Texas, Austin,
and Ririko Horvath, Larry May, and Rahul Tikekar, Research, Analysis, and Statistics,
Internal Revenue Service*

Introduction

The tax law allows flow-through tax treatment for certain legal entities so that income is subject to tax only once—at the partner or shareholder level. The information return that flow-through entities file includes a Schedule K-1 that shows income, deductions, credits, and other items that are allocated to the owners. Flow-through treatment introduces complexity in auditing tax returns as flow-through entities can be associated with many owners and vice versa. A flow-through entity can also be associated with many other flow-through entities. Additionally, the financial flows can vary across owners—both in type and magnitude. For example, owners can receive different types of income and deductions from a flow-through entity, and each owner may not receive the same proportion of the income and deductions. Due to this complexity, there is a need to represent the associations between different types of entities and the related financial flows in a form that is quantifiable and that can be used to evaluate these associations for tax compliance risk. This is especially important considering that more than 20 million Schedule K-1s are issued every year.

In its simplest form, a tax structure that includes flow-through entities is a network where entities can be represented as nodes and the linkages between entities can be represented as edges. Further, the financial flow associated with the linkages can represent the strength or weight of the link. Thus, social network analysis (SNA) is a potential tool that can be used to represent the economic structures resulting from the use of flow-through entities. SNA has been successfully used to represent the complex associations and flows across heterogeneous entities such as individuals, products, and firms. Many commercial platforms such as Facebook, Twitter, and Google+ use SNA to capture the dynamic linkages between individuals and associate individual behavior and preferences with their network characteristics. These platforms use these insights to provide targeted information and services to these individuals. Prior research has also investigated interconnections between user networks and product networks. Similarly, SNA has been used to represent industry structures and the associated outcomes.

In this paper, we investigate how SNA can be applied to characterize the complex associations between different entities within a flow-through structure and develop measures to quantify these associations.¹ To achieve this, we consider two different samples of Schedule K-1 data obtained from the Internal Revenue Service (IRS) yK1 database. We also consider different types of linkages occurring within enterprises, namely K-1 linkages, primary-secondary associations (i.e., spousal linkages), and parent-subsidiary linkages. For our analysis, we first examine existing enterprises and develop SNA measures at both the enterprise and node levels. Next, we construct graphs representing the different enterprises in our two yK1 samples and calculate SNA measures. We also investigate the potential application of these graphs to identify economically important nodes and unusual enterprises.

Our investigation shows that SNA can be used to represent the tax structures associated with flow-through entities. This includes the ability to capture different types of nodes, different types of linkages across these nodes, and the ability to represent various types of financial flows associated with the entities. We also illustrate how network measures can be used to characterize different enterprises and compare these enterprises. More specifically, we show how SNA measures can be used to determine if enterprises conform to the expected network structure and to flag any exceptions. Finally, we demonstrate how a combination of network and node level measures can be used to identify economically important nodes in an enterprise.

¹ In future work, we plan to investigate whether these SNA measures are predictive of tax noncompliant behavior.

Prior Literature

Prior academic work documents that certain firm characteristics, such as magnitude of book-tax differences, firm size, industry, and multinationality, are associated with corporate tax noncompliance (Mills (1998); Hanlon, Mills, and Slemrod (2007)). There is also evidence that greater organizational complexity and financial complexity are associated with higher levels of corporate tax avoidance (Wagener and Watrin (2013); Balakrishnan, Blouin, and Guay (2012)). However, these prior studies are limited to drawing conclusions about consolidated corporate entities where publicly filed financial statement data are available. In general, academic tax research has focused on corporate tax issues due to these data constraints.

With respect to flow-through entities, a separate stream of academic literature examines the choice of overall business structure given the tax and nontax costs and benefits (e.g., Guenther (1992); Ayers, Cloyd, and Robinson (1996); Gordon and MacKie-Mason (1994); MacKie-Mason and Gordon (1997)). For example, Ayers, *et al.* (1996) find that in choosing between corporate and noncorporate structures, business risk, number of owners, firm size, and firm age influence the choice of organizational form. However, again, these studies are limited to examining the organizational form decisions of top-level entities and cannot provide evidence regarding the use of flow-through entities embedded within corporate structures. Recent work on the use of special purpose entities, which include LLCs, LLPs, trusts, and other entities, is a first step in addressing how enterprises use flow-through entities within their business structure (Feng, Gramlich, and Gupta (2009); Demere, Donohoe, and Lisowsky (2015)). Nonetheless, without more detailed data, these studies can draw only high-level conclusions.

In this paper, we have access to a unique dataset collected by the IRS, which gives us visibility into the underlying organizational structure of business enterprises. Given the complexity of the data, which prior studies could not observe, we propose that SNA techniques can be useful in quantifying the many dimensions of complexity. For example, not only are we interested in measuring the number and types of entities within an enterprise, we also aim to measure and quantify the shape of enterprise structures, how entities are related, and the magnitude of these relationships. Our study contributes to the literature by proposing a new methodology to measure and quantify business structures.

Data

Our unit of analysis is an enterprise that contains two or more entities and includes at least one flow-through entity. These enterprises are defined by IRS using a 50-percent ownership rule. Specifically, a flow-through entity is considered to be part of an enterprise only if the taxpaying entities associated with the enterprise have—directly or indirectly—at least 50-percent ownership of the flow-through entity.

We consider two different samples of enterprises to conduct our proposed social network analysis. The first sample is based on an intersection of enterprises associated with entities that also appear in the proposed deficiency database in Tax Year 2009. The entities in the proposed deficiency database file Form 1120. This database flags entities that have deficiencies in their tax filings and also reports the proposed deficiency amounts. Thus, SNA measures of this sample of enterprises can potentially allow us in future work to determine the association between enterprise network characteristics and tax noncompliance. We also consider another sample, which represents all enterprises with flow-through entities. For our second sample, we randomly select 5,000 enterprises from Tax Year 2009 that have between 5 and 15 nodes, and we consider all entities associated with these enterprises. The objective of selecting this second sample is to compare the network characteristics of the proposed deficiency sample with a random sample of entities. Table 1 provides a summary of the two data samples. We extract the information about linkages between the sample entities. We consider parent-subsidiary, primary-secondary, and K-1 linkages. We also extract the financial flows associated with each link, such as gains, losses, income, interest, capital gains, rent, real estate income, etc. If there are multiple entries associated with a particular type of financial flow for a given link, we sum these values.

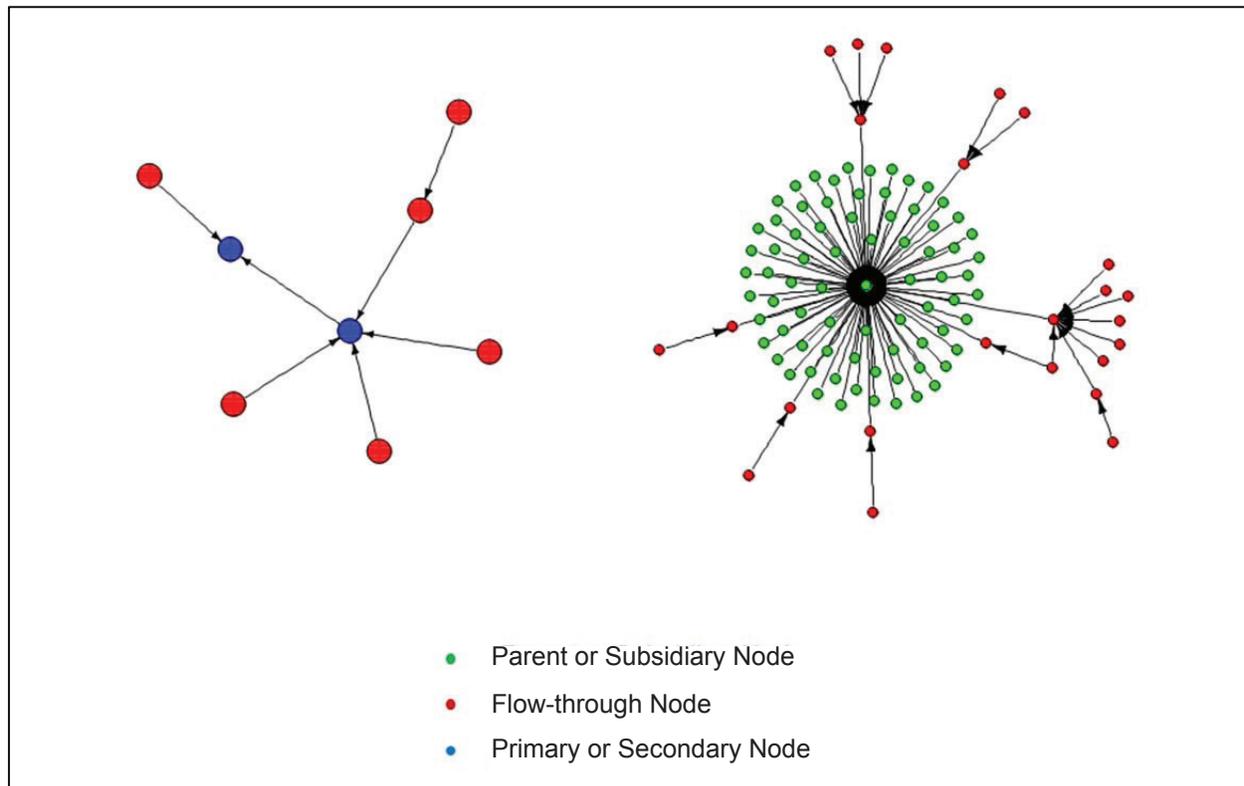
TABLE 1. Data Samples

	Sample Based on Proposed Deficiency Database	Random Sample
Year	2009	2009
Number of enterprises	5,913	5,000
Entities	107,638	31,884
K-1 links	411,644	28,210
Parent-Sub links	75,832	1,225
Primary-Secondary links	55	2,590

Social Network Analysis

We use the igraph library in R CRAN for building the network and calculating SNA measures. We treat each enterprise as an independent network and create separate graphs for each enterprise. Each enterprise graph is both a directed and weighted network. In a directed network, transactions always flow from the payer to the payee. The weights represent different types of flows where each weight can vary in magnitude. Each type of flow (income, gain, loss, dividend, interest, etc.) can have a separate weight based on the magnitude. We tag nodes based on the type of connections. Figure 1 shows two sample enterprises. We create enterprise-level SNA measures that represent the entire network associated with the enterprise. We also create node-level SNA measures that are associated with individual entities within each enterprise. These measures are described in the next section.

FIGURE 1. Sample Networks Representing Enterprises



Network Measures

Our primary objective is to determine if the network obtained by the use of flow-through entities and the associated economic flows represent tax-noncompliance risk. Recent studies suggest that complex enterprises are associated with higher levels of tax avoidance than more simple structures (Balakrishnan, *et al.* (2012), Wagener and Watrin (2013)). This suggests that complexity of the economic structure associated with the use of flow-through entities could potentially reveal tax avoidance (and perhaps evasion) behavior. To represent this complexity, we use standard network theory to develop different measures to characterize the network associated with individual entities and the enterprise. Our measures represent different combinations of the attributes of the entities involved in an enterprise and the economic flows associated with these nodes.

Network-Level Measures

Studies show that flow-through entities facilitate multistate tax avoidance and have been widely used as special purpose vehicles (Fox and Luna (2005), Feng, *et al.* (2009)). Therefore, the characteristics and distribution of flow-through entities in an enterprise may represent risk for tax noncompliance. To capture this, we include network-level measures based on the characteristics of the flow-through entities. These are described below.

1. Density

A denser network with a greater number of connections between nodes is more complex than a network with an equal number of nodes and fewer connections. Thus, we may expect a network or an enterprise with high density to have higher compliance risk than other enterprises.

Density can be defined as the ratio of the number of links present in the network to the maximum number that are logically possible, given the size of the network. It can be captured as the following ratio:

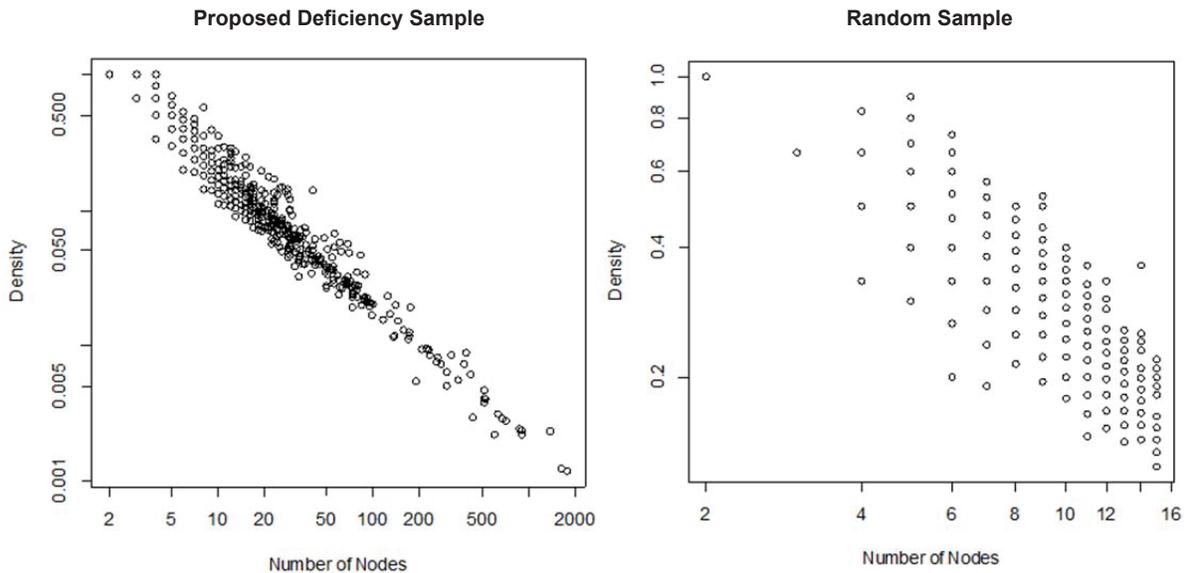
$$\frac{\text{Total Number of Links}}{\frac{1}{2}n(n-1)}$$

where

n = number of nodes within an enterprise.

Note that enterprises with fewer nodes will have higher density. In that case, to compare across enterprises of different sizes, this measure can be normalized by size. We can use either the number of nodes or total enterprise assets to represent the size. Figure 2 compares the density values associated with enterprises in our two data samples as a function of the number of nodes. (Refer to the descriptions of these two samples above.)

FIGURE 2. Enterprise Density as a Function of the Number of Nodes



2. Diversity of Nodes

A network including different types of nodes is more complex and may reflect greater tax planning than a homogenous network. Thus, we may expect a network or an enterprise with high diversity of nodes to have higher compliance risk than other enterprises. Diversity of nodes can be defined as a degree of concentration where nodes are of different types. It can be captured as the Simpson index or Herfindahl index. The measure equals the probability that two entities taken at random from the network represent the same type. It equals:

$$\sum_i p_i^2$$

where p_i represents the proportion of node of type i . Lower values of the index indicate lower concentration and thus greater diversity.

3. Loss Nodes

A network with an abnormally high proportion of flow-through entities incurring losses may reflect greater tax planning than a network with a smaller proportion of flow-through loss nodes. Economically, we would not expect an

enterprise to continue operating unprofitable nodes. Thus, the presence of a high number of loss nodes could be indicative of noncompliance as the losses would provide tax savings to taxpaying entities.

Loss nodes can be defined as the proportion of flow-through entities incurring losses within the network to the total number of nodes. It can be captured as the following ratio:

$$\frac{\text{Number of Flow-through Loss Nodes}}{\text{Total Number of Nodes}} .$$

Note that larger enterprises are expected to have lower proportions of loss nodes due to diversification.

4. External Degree Centrality

The interaction between enterprises, perhaps through joint ventures or minority partnership interests, is more complex than a self-contained network. Hence, we may expect a network with a high number of external linkages to have higher compliance risk than other networks.

External degree centrality can be defined as the proportion of links within a network that are connected to entities external to the immediate network. It can be captured as the following ratio:

$$\frac{\text{Total Number of External Links Associated with Other Enterprises}}{\text{Total Number of Links}} .$$

5. Graph Centralization

Graph centralization represents the variation in the centrality scores of the nodes in a network. Centrality of a node reflects its importance locally (degree) or relative to the rest of the network (closeness, betweenness, etc.). A highly centralized graph represents a structure where only a few nodes are the focus of the economic activity.

Centralization can be expressed as:

$$\frac{\sum_i [C_D(\text{max}) - C_D(i)]}{(n-1)(n-2)}$$

where $C_D(i)$ is the centrality of node i (see below) and n is the number of nodes within an enterprise.

Node-Level Measures

The purpose of developing these measures is to highlight economically important entities within an enterprise. These measures can be aggregated further across all the entities within an enterprise to come up with a composite score at the enterprise level.

1. Degree Centrality (Standardized)

Nodes with high degree centrality are connected to a greater number of entities within the network structure. This represents a higher level of complexity than nodes with few connections. Thus, we may expect a node with high degree centrality to have higher compliance risk than other nodes.

Degree centrality can be defined as the number of linkages present at each node. It can be captured as:

$$\frac{\text{Number of Links per Node}}{n-1}$$

where n = number of nodes within an enterprise.

A node with a large asset balance is expected to have high degree centrality. Thus, to compare across nodes, we normalize this measure by total node-level assets.

2. Weighted Degree Centrality

Nodes with high degree centrality adjusted for their level of activity are more likely to be economically important entities within the network structure. We may expect a node with greater activity to have higher compliance risk than other nodes.

Weighted degree centrality can be defined as the number of linkages present at each node weighted by various types of economic flows. Our primary economic flows of interest are profits and losses. We can also weight this measure by interest and royalty payments as these flows are associated with tax-planning strategies using intercompany loans and intellectual property transfers. Weighted degree centrality can be captured as:

$$\text{Number of Links per Node}^{1-\alpha} \times \text{Node Strength}^\alpha$$

where

$$\text{Node Strength for a node } i = \sum_{j \neq i}^n w_{ij}$$

w_{ij} is the weight representing the economic flow between two nodes i and j , and

α is a tuning parameter determining the relative importance of number of links compared to the weight of links.

The tuning parameter can be determined by estimating the relative impact of each attribute on the desired output. Also note that the weight of a link is equal to the magnitude of the economic flow.

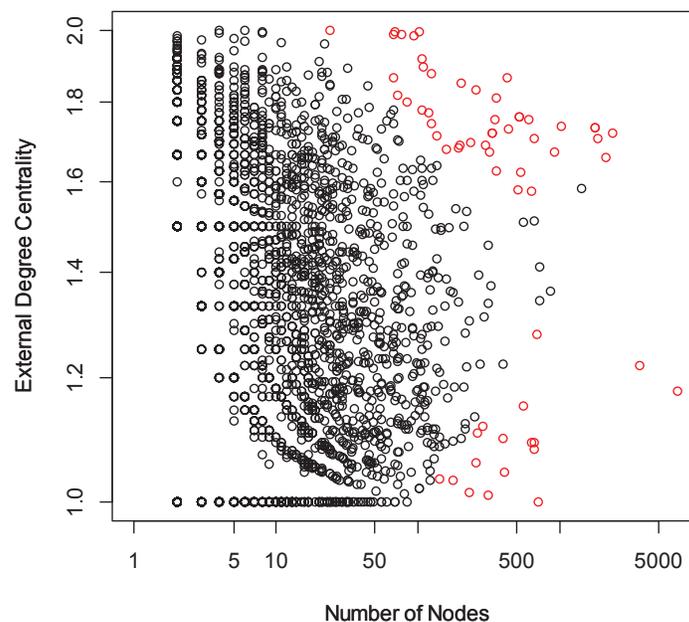
Other node-level measures are explained in the Appendix.

Application of SNA

Outlier Analysis

One of the applications of SNA is to create common measures to describe enterprises/nodes and to use these measures to find exceptions. Such exceptions in our context can potentially point to noncompliant tax behavior. For example, we can expect that the external degree centrality of an enterprise decreases as the number of nodes associated with the enterprise increases. Large deviations from the expected value of this measure can be flagged as exception. To determine such exceptions or outliers, we can carry out simple regression analysis to model expected behavior. Outliers can be identified as deviations from the fitted values obtained from the regression coefficient. Figure 3 shows the outlier enterprises (marked in red) using Cook's Distance, a commonly used measure to identify outliers.

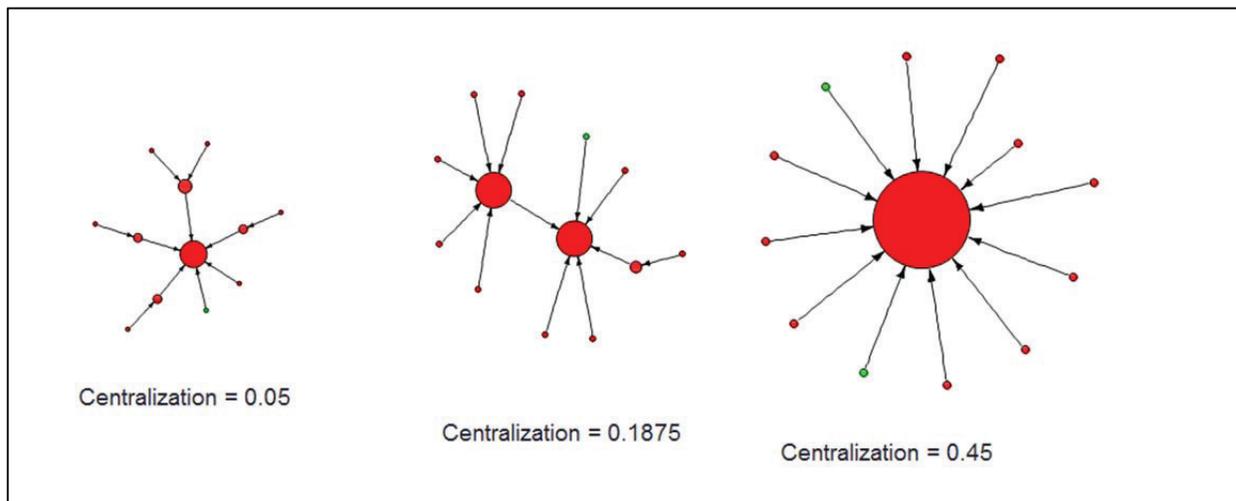
FIGURE 3. External Degree Centrality



Identifying Economically Important Nodes

Graph centralization and degree centrality of the individual nodes can be combined to determine economically important nodes. A graph with a high degree of centralization indicates that certain nodes are dominant in the enterprise. These dominant nodes can be identified using their degree centrality values. A dominant node is expected to have a higher degree centrality value. Figure 4 shows sample graphs for different enterprises, which vary in their centralization values. The node size in each graph is proportional to the degree centrality of the node. The enterprise with a high centralization value also has nodes with a much higher degree centrality value as compared to the other nodes. In these graphs, we have used standard degree centrality values. However, the same analysis can be repeated using weighted degree centrality, where one can use any of the previously defined weights.

FIGURE 4. Enterprises With Different Centralization Values



Conclusion and Future Work

We investigate how SNA techniques can be used to analyze tax structures that include flow-through entities. We show that SNA can be a useful approach to characterize these structures. It allows us to represent diverse enterprises consisting of different types of taxpaying entities and flow-through entities. Further, we can capture different types of financial flows associated with these entities. SNA provides measures at the enterprise level and the node level within an enterprise. This allows efficient comparison of these entities across several different metrics. SNA can be used further to identify exceptions both at the enterprise level and at the node level within an enterprise.

While the current work illustrates the potential of using SNA techniques to analyze flow-through entities, there are several additional avenues for future work to establish the use of SNA in predicting tax noncompliance and to operationalize these measures for implementation. These include:

- Measure Validation:* Future work should conduct empirical analysis using existing noncompliance data for enterprises and identify relevant SNA measures that are indicators of noncompliance.
- Exception or Outlier Identification:* Future work should focus on defining robust measures for identifying outliers or exceptions. This involves establishing the correct association between the size of the enterprise and the SNA measures and investigating different metrics to identify outliers. Additionally, data-mining approaches can be explored to conduct the outlier analysis.
- Measure Refinement:* Future work should also dig deeper into the definition of the proposed measures. This includes validation of the measures using a training dataset. Additionally, there is significant overlap in the proposed measures. Thus, future work should investigate composite measures that can be used to characterize enterprises.
- Enterprise Definition:* Currently, we rely on the enterprise definition based on the 50-percent rule. Future work should investigate network structures without imposing this rule and conduct the outlier analysis.

- e) *Multi-Year Analysis*: Currently, we focus on a single year to establish the use of SNA. Future work should validate the SNA measures across multiple years.

Besides establishing the role of SNA to characterize tax structures, another potential avenue of research is using SNA to evaluate these structures over time. Taxpaying entities may alter their structure in response to tax law changes or audits. In that case, the measure thresholds used to identify exceptions should change over time. Future work should also incorporate this dynamic aspect in the definition of SNA measures.

References

- Ayers, B., C. Cloyd, and J. Robinson. 1996. "Organizational form and taxes: An empirical analysis of small businesses." *Journal of the American Taxation Association* 18(1): 49–67.
- Balakrishnan, K., J. Blouin, and W. Guay. 2012. "Does tax aggressiveness reduce corporate transparency?" Working paper, University of Pennsylvania.
- Demere, P., M. Donohoe, and P. Lisowsky. 2015. "The economic effects of special purpose entities on corporate tax avoidance." Working paper, University of Illinois.
- Feng, M., J. D. Gramlich, and S. Gupta. 2009. "Special purpose vehicles: Empirical evidence on determinants and earnings management." *The Accounting Review* 84 (6): 1833–1876.
- Fox, W. F., and L. Luna. 2005. "Do limited liability companies explain declining State corporate tax revenues?" *Public Finance Review* 33 (6): 690–720.
- Gordon, R., and J. MacKie-Mason. 1994. "Tax distortions to the choice of organizational form." *Journal of Public Economics* 55 (2): 279–306.
- Guenther, D. 1992. "Taxes and organizational form: A comparison of corporations and master limited partnerships." *The Accounting Review* 67 (1): 17–45.
- Hanlon, M., L. Mills, and J. Slemrod. 2007. "An empirical examination of big business tax noncompliance." In *Taxing Corporate Income in the 21st Century*, edited by A. Auerbach, J. Hines, Jr., and J. Slemrod, 171–210. Cambridge: Cambridge University Press.
- MacKie-Mason, J., and R. Gordon. 1997. "How much do taxes discourage incorporation?" *Journal of Finance* 52 (2): 477–505.
- Mills, L. 1998. "Book-tax differences and Internal Revenue Service adjustments." *Journal of Accounting Research* 36 (2): 343–356.
- Wagener, T., and C. Watrin. 2013. "The relevance of complex group structures for income shifting and investors' valuation of tax avoidance." Working paper, University of Münster.

Appendix

Additional Node-Level Measures

1. Closeness Centrality (Standardized)

Nodes with low levels of closeness are likely entities within a multitier network structure, which is more complex than a flat structure. Hence, we may expect a top-level node with lower closeness to have higher compliance risk than other nodes. We may also use the closeness centrality measure of the top-level node to represent the overall closeness of the network or enterprise.

Closeness centrality can be defined as the inverse of the distance between a node and every other node in the network, where distance is measured as the number of links in the shortest path from one node to another. It can be captured as:

$$\left[\frac{\sum_{j=1}^n \text{Distance}(i, j)}{n-1} \right]^{-1}$$

where

n =number of nodes within an enterprise

i, j =node i , node j , etc.

Large enterprises can have lower closeness. To compare across enterprises we can normalize this measure by total enterprise assets.

2. Profit and Loss Asymmetry

Skewed allocations of profit and loss can be an indication of tax planning. For example, allocating a partner a high percentage of flow-through losses but a low percentage of flow-through profits does not appear to be an economically rational allocation rule. Thus, we may expect a node with a high level of profit and loss asymmetry to have higher compliance risk than other nodes.

Profit and loss asymmetry can be defined as the disparity between an entity's share of profits and share of losses. It can be captured using the "Node Strength" where:

$$\text{Node Strength for a node } i = \sum_{j \neq i}^n w_{ij}$$

where w_{ij} is the weight representing the economic flow between two nodes i and j .

In this case, the weight of a link is equal to the absolute value of the difference between percentage allocation of profits and percentage allocation of losses at each link. Use of absolute differences can allow us to capture all deviations.

3. Net Flows Asymmetry

Noncompliant behavior could include unusual patterns of flows between entities, particularly where net flows are not economically rational or lack economic substance. Thus, we may expect a flow-through node with a high level of net flows asymmetry to have higher compliance risk than other flow-through nodes.

Net flows asymmetry captures where inflows and outflows to and from flow-through nodes are mismatched in terms of sign (e.g., inflows are all positive but outflows are all negative). It can be calculated as:

$$\text{Node Strength}_{\text{inflows}} - \text{Node Strength}_{\text{outflows}}$$

4. *Character Asymmetry*

Noncompliant behavior could also involve unusual patterns of types of flows between entities, such as disproportionate flows of tax-preferred income items and deductions against ordinary income. Thus, we may expect a flow-through node with a high level of character asymmetry to have higher compliance risk than other flow-through nodes.

Character asymmetry captures where inflows and outflows to and from flow-through nodes are mismatched in terms of the character of the income or loss (e.g., tax preferred vs. ordinary items). It can be calculated as:

$$\frac{\textit{Tax-Preferred Inflows (Dividends, Capital Gains)}}{\textit{Total Inflows}} - \frac{\textit{Tax-Preferred Outflows}}{\textit{Total Outflows}}$$



Taxpayer Responses to Rules and Enforcement

Ayers ♦ Seidman ♦ Towery

Boynton ♦ DeFilippes ♦ Legel ♦ Rupert

Datta ♦ Orlett ♦ Turk

Taxpayer Behavior Under Audit Certainty

*Benjamin C. Ayers and Erin M. Towery, University of Georgia, and
Jeri K. Seidman, University of Texas at Austin¹*

Over the past several decades, both theoretical and empirical studies have documented that the risk of tax audit (examination by the tax authority) affects taxpayer behavior. These studies generally predict that taxpayers enter fewer uncertain tax positions to reduce their probability of audit because tax uncertainty increases the probability of audit. However, some taxpayers face a *certain* annual audit, meaning that the level of tax uncertainty does not affect their probability of tax audit. Understanding how audit certainty affects taxpayer behavior is important because many of the largest firms in the U.S. face certain audit every year, and the Internal Revenue Service (IRS) invests a sizeable portion of its resources in these efforts. Audit certainty therefore impacts an economically significant proportion of the U.S. economy, and the IRS commitment to a program of audit certainty is an economically important resource allocation decision. We know little about the effect of audit certainty on taxpayer behavior because taxpayers are not required to disclose publicly whether they face certain audit.

In this study, we overcome this data limitation using a confidential dataset of corporate taxpayers where the risk of audit is 100 percent—firms in the IRS Coordinated Industry Case (CIC) program. The IRS implemented the CIC program (formerly the Coordinated Examination Program) in the 1960s in response to the growing complexity of U.S. business operations. For CIC firms, a team from the IRS's Large Business and International (LB&I) group spends a substantial amount of time in the taxpayer's primary place of business throughout the year. The IRS team consists of the examination team manager, field agents, industry specialists, and subject-matter experts.²

The effect of audit certainty on taxpayer behavior is not clear *ex ante*. On the one hand, taxpayers could have less incentive to engage in tax avoidance (or evasion) if the increased audit probability decreases the expected benefit of tax avoidance such that a subset of tax positions are no longer value-creating. This would be consistent with the negative relation between audit risk and tax avoidance documented at lower points on the audit probability spectrum (Hoopes, Mescall, and Pittman (2012)).

On the other hand, Mills and Sansing (2000) suggest that certainty of audit could increase the incentive to engage in tax avoidance. Specifically, because the IRS will audit the firm regardless of the signals provided in the financial statements, certain audit firms no longer have incentive to reduce the difference between book and taxable income. This intuition is consistent with results presented in Slemrod, Blumenthal, and Christian (2001), which reports results of a 1995 experiment by the Minnesota Department of Revenue under which a random sample of individual taxpayers were told that the returns that they were about to file would be closely examined. Relative to the sample not told this, the high-income members of the "audit certain" sample significantly decreased their reported tax liability. The authors conjecture that these individual taxpayers claim more tax benefits to create a more aggressive starting point for negotiations with the goal of minimizing tax liability, assuming the audit will not detect and punish all tax avoidance. Further, the authors postulate that this effect is observed in high-income taxpayers, but not low- or middle-income taxpayers, because the high-income taxpayers believe that the final outcome of the certain audit is more manipulable, likely in part due to their ability to hire professional assistance. This logic likely applies to the corporate taxpayers we study, as firms assigned to the CIC program tend to be large and/or have complex operations, both of which are likely correlated with the likelihood of professional tax assistance. However, the corporate taxpayers in our sample have financial reporting obligations that individuals do not, which could cause these two types of taxpayers to have different tax avoidance preferences. In sum, how tax certainty affects taxpayer behavior is an empirical question.

¹ The authors appreciate helpful comments from Christine Cheng (discussant), Danielle Higgins (discussant), and workshop participants at the University of Houston, College of William & Mary, University of Virginia McIntire School of Commerce, University of Kansas, the 2014 American Accounting Association Annual Meeting, and the 2015 IRS TPC Research Conference. We also thank John Miller and Barbara Hecimovich for providing information about the IRS CIC program.

² CIC audit teams generally provide more in-depth audits than traditional IRS audits. For example, subject-matter experts in areas such as engineering, excise taxes, and employment are included in the list of specialists assigned to a CIC audit team.

Before testing our research question, we first analyze the determinants of assignment to the CIC program using CIC selection factors outlined in the Internal Revenue Manual. CIC assignment is based on a point system involving seven main criteria: (i) gross assets; (ii) gross receipts; (iii) operating entities; (iv) number of industries; (v) total foreign assets; (vi) related transactions; and (vii) foreign taxes paid. Each criterion has a point value, and a firm is assigned to the CIC program if its total point value is greater than or equal to 12. Firms with a point value less than 12 can also be assigned to the CIC program if they are sufficiently complex to warrant certain audit.

Though other research (Mills (1998); Hanlon, Mills, and Slemrod (2007)) cite size and complexity as determinants of CIC assignment, these statements are based on the Internal Revenue Manual's listed factors rather than on empirical tests. Thus, our analysis serves two purposes: 1) to shed light on whether CIC program assignment is based on factors outlined by the IRS and/or factors associated with tax avoidance, and 2) to provide researchers without access to CIC-assignment data a model of audit certainty. We find that many of the selection factors are positively associated with assignment into the CIC program, with gross receipts being the most significant size determinant and the number of geographic segments being the most significant complexity determinant. When we include factors known to affect firms' incentives or ability to avoid taxes (such as research and development expenses, excess stock option deductions, and net operating loss carryforwards) as potential determinants, we find that some of the factors are significantly associated with CIC assignment. However, their inclusion does not dramatically improve the fit of the model. These results collectively suggest that although inclusion in the CIC program is associated with firms' incentives or ability to avoid taxes, the CIC assignment decision is primarily based on firm size and complexity.

Next, we study the effect of audit certainty on taxpayers' initial filing liabilities and tax reserves. We use the taxpayer's initial Federal filing liability rate as our primary variable to test whether audit certainty has a deterrent effect on tax avoidance behavior. The advantage of the taxpayer's initial Federal liability rate is that it captures initial tax payments to the tax authority. We use the taxpayer's cash effective tax rate (ETR) as an alternative, publicly available proxy for tax payments to tax authorities. Reserves for uncertain tax positions reported in financial statements proxy for managers' expectations of future tax payments associated with current tax return positions.

We test our prediction regarding the effect of certain audit on taxpayer behavior using both a levels approach and a changes approach. We implement our levels analysis using a pooled sample from 2000 to 2011 of firms assigned to the CIC program and firms not assigned to the CIC program. We find that firms that are assigned to the CIC program do not report higher Federal filing liability rates or cash effective tax rates than firms who are not assigned to the program.

To implement our changes analysis, we first identify 405 corporate taxpayers that are assigned to the CIC program for the first time between 2000 and 2011 ("newly assigned firms").³ We then construct two samples of propensity-matched control firms—(i) firms not assigned to the program ("nonassigned firms"), and (ii) firms assigned to the program for at least the prior four years that remain assigned to the program ("long-assigned firms"). The matched sample design allows us to not only compare the tax behavior of the firm to itself before and after the change in its CIC program status, but to also compare its tax behavior with the tax behavior of a firm that does not experience a change in CIC status. We find that, post-assignment, neither the Federal filing liability rates nor the cash effective tax rates of newly assigned firms are statistically different than those of the matched sample of nonassigned firms. Further, post-assignment, neither of the tax payment rates of newly assigned firms is statistically different than the tax payment rates of the matched sample of long-assigned firms. Thus, our results suggest audit certainty does not result in firms decreasing tax avoidance upon program assignment.

However, we estimate that newly assigned firms report higher reserves for current-year tax positions relative to both nonassigned and long-assigned firms, suggesting that audit certainty does impact financial reporting for income taxes. Our result that the initial tax liability does not change for newly assigned firms suggests that the increased reserves do not represent an increase in aggressive tax avoidance. More plausible explanations include: (i) firms systematically underestimated their likelihood of sustaining a position prior to CIC assignment and subsequently update their expectations based on "learning" in the audit process, and/or (ii) firms incorporated audit likelihood in their determination of reserves prior to CIC assignment (contrary to the U.S. GAAP requirement that firms assume audit certainty with respect to each uncertain position).

³ While the probability of firm audit is 100 percent in the CIC program, the probability of audit for any particular transaction remains less than 100 percent. Thus, taxpayer behavior may continue to affect the audit risk of any particular transaction. Even so, assignment to the CIC program represents a significant positive shock to the audit probability of the firm and thus the audit probability of any particular item.

Our study expands the academic literature in two important ways. First, our model of CIC determinants provides researchers with a better proxy for the audit risk of large, publicly traded companies. Prior studies measure CIC participation as firms with at least \$250 million in assets (e.g., El Ghouli, Guedhami, and Pittman (2011); Hoopes, *et al.* (2012)). We report that only 19.5 percent of firms with assets greater than \$250M are assigned to the CIC program, suggesting that this commonly used proxy is quite weak. Second, to our knowledge, our study is the first to analyze corporate taxpayer behavior under audit certainty. In doing so, we further our understanding of the strategic game between the taxpayer and the tax authority. Though the strategic tax model does not postulate a corner solution, our results suggest that audit certainty affects taxpayer behavior in ways inconsistent with the strategic tax model and thus suggests the need for a more complete model.

Our study also informs tax authorities. Understanding how audit risk affects taxpayer behavior is important to the IRS as they design and implement new audit approaches. Per a discussion between one of the authors and the IRS, CIC audits consume a substantial portion of IRS's Large Business and International (LB&I) audit resources. Whether and how firms alter behavior within the CIC program informs the cost-benefit assessment of the program. We do not find a decrease in tax avoidance within the CIC program. However, we do find an increase in unrecognized tax benefits, which suggests that audit certainty increases managers' expectations of future tax payments but that the IRS must exert audit effort to generate significant benefits from the CIC program (i.e., via disallowing tax positions). Consistent with this interpretation, we document that IRS settlements are larger for firms in the CIC program, suggesting that the CIC program generates Treasury revenue primarily via an enforcement effect.

References

- El Ghouli, S., O. Guedhami, and J. Pittman. 2011. "The Role of IRS Monitoring in Equity Pricing in Public Firms." *Contemporary Accounting Research* 28(2): 643–674.
- Hanlon, M., L. Mills, and J. Slemrod. 2007. "An Empirical Examination of Corporate Tax Noncompliance." In: Auerbach, A., Hines, J., Slemrod, J. (Eds.), *Taxing Corporate Income in the 21st Century*. Cambridge University Press, New York, 171–210.
- Hoopes, J. L., D. Mescall, and J. A. Pittman. 2012. "Do IRS Audits Deter Corporate Tax Avoidance?" *The Accounting Review* 87(5): 1603–1639.
- Mills, L., 1998. "Book-tax Differences and Internal Revenue Service Adjustments." *Journal of Accounting Research* 36: 343–356.
- Mills, L. F., and R. C. Sansing. 2000. "Strategic Tax and Financial Reporting Decisions: Theory and Evidence." *Contemporary Accounting Research* 17(1): 85–106.
- Slemrod, J., M. Blumenthal, and C. Christian. 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79(3): 455–483.

2011–2012 Schedule M-3 Profiles of Schedule UTP Filers by IRC Section Cited*

*Charles Boynton, Ellen Legel, and Lisa Rupert, Large Business and International Division, Internal Revenue Service, and Portia DeFilippes, Office of Tax Analysis, U.S. Department of the Treasury*¹

Part I. Schedules M-3 and UTP Background

1. Schedule M-3 Overview

Taxpayers prepare corporate and partnership tax returns by adjusting amounts from their financial statements or books and records. The goal of the Schedule M-3 reconciliation is to increase taxpayer transparency to the Internal Revenue Service (IRS) with respect to the book-to-tax differences (BTD) resulting from adjustments made to financial statements or books and records in preparing the tax return and to assist the IRS in selecting returns and issues for audit where tax compliance risk is present and in not selecting returns and issues where such risk is not present.

Schedule M-3 was first introduced in 2004 for U.S. corporations with total assets of \$10 million or more filing U.S. income tax return Form 1120. It replaced four decades of use of the less structured Schedule M-1 for these corporations for the required reconciliation of financial statement income to tax income.²

A Treasury report in 1999 and Treasury testimony in 2000 by Assistant Secretary (Tax Policy) Jonathan Talisman viewed the 1990s widening difference between the sum of corporate financial statement income (book income) and Federal income tax expense reported on Form 1120, Schedule M-1, lines 1 and 2, and tax income reported on Form 1120, page 1, line 28, as a possible indicator of corporate tax shelter activity, but also noted the difficulty in interpreting Schedule M-1 BTD data.³

Mills and Plesko (2003) proposed a redesign of Schedule M-1 to increase the transparency of the corporate book-to-tax reconciliation and to improve data interpretability.⁴ The Mills and Plesko (2003) Schedule M-1 redesign recommendations are largely reflected in Schedule M-3, particularly in Part I.⁵

Schedule M-3, Part I, is important and unique in tax reporting in that it lists the adjustments made to worldwide consolidated income in the parent corporation's financial statements to determine the book income

* First published in *Tax Notes* Volume 149, No. 2, (October 12, 2015), page 249. Reprinted with permission. Prepared for the 2015 IRS-TPC Research Conference.

¹ The opinions expressed are those of the authors and do not necessarily represent positions of the U.S. Department of the Treasury or the Internal Revenue Service.

² The current paper repeats certain material from Boynton, DeFilippes, and Legel (2005, 2006a, 2006b, and 2008), Boynton, DeFilippes, Legel, and Reum (2011 and 2014), Boynton, DeFilippes, Legel, and Rupert (2014), Boynton and Wilson (2006), and Boynton and Livingston (2010) used with permission. Our tax return table values may not add due to rounding. The SOI corporate data file for year t includes all tax years ending between July of Calendar Year t and June of Calendar Year $t+1$. Effective for tax years ending on or after December 31, 2004, Schedule M-3 replaced Schedule M-1 for corporations filing Form 1120 and reporting total assets of \$10 million or more on Form 1120, Schedule L. Effective for tax years ending on or after December 31, 2006, for corporations with total assets of \$10 million or more, Schedule M-3 applies to Form 1120S for S corporations, to Form 1120-C for cooperative associations, and to Form 1120-L for life insurance companies and Form 1120-PC for property and casualty insurance companies. Effective for tax years ending on or after December 31, 2006, Schedule M-3 also applies to Forms 1065 and 1065-B for partnerships with total assets of \$10 million or more and to certain other partnerships. Effective for tax years ending on or after December 31, 2007, a special Schedule M-3 applies to Form 1120-F for foreign corporations with effectively connected U.S. income and total assets of \$10 million or more. Schedule M-1 continues to apply to Form 1120-RIC for regulated investment companies, to Form 1120-REIT for real estate investment trusts, and to all corporations with total assets of less than \$10 million. Effective for tax years ending December 31, 2014, and later, corporations and partnerships with \$10 million or more in assets but less than \$50 million in assets and those partnerships with less than \$10 million in assets required to file Schedule M-3 would be permitted to file Schedule M-3, Part I, and to file Schedule M-1 in place of Schedule M-3, Parts II and III, if they so choose.

³ See U.S. Department of the Treasury (1999) and Talisman (2000). See also Mills (1998) cited by Treasury (1999, page 32, note 118).

⁴ See Mills and Plesko (2003) for the proposed redesign of Schedule M-1.

⁵ For a discussion of the development of Schedule M-3, see Boynton and Mills (2004).

of the includible corporations in the tax return.⁶ We also use Part I data to identify each corporation financial statement type as Securities and Exchange Commission (SEC) 10K/Public, Audited, or Unaudited.⁷

Schedule M-3, Parts II and III, are a more structured listing of BTD than Schedule M-1 and specify a number of fixed categories as well as two “other with difference” categories. The fixed categories are machine readable. The book income and tax income amounts generating the BTD are listed as well as the BTD and the name for the line.

On Schedule M-3, Parts II and III, BTD are characterized as temporary or permanent. Temporary differences are items of income or expense that are recognized for both financial and tax reporting, but appear in different time periods. Permanent differences are items of income or expense that are recognized for either financial or tax reporting, but not both.⁸

Parts II and III contain four columns. Column (a) represents financial statement (book) income or expense amounts using the financial statement source determined in Part I. Column (d) represents amounts as shown on the tax return. The BTD between the amount shown in column (a) and the amount shown in column (d) is reported either as a temporary difference amount in column (b) or as a permanent difference amount in column (c).

*Note that on Schedule M-3, a negative total BTD adjustment occurs if tax income is below book income. Further note that in our study we conform the sign of Part III data to agree with Part II so that a negative book income or tax income item always reduces total book income or tax income and a negative BTD reduces tax income.*⁹

We impose certain minimum reconciliation requirements on the returns included in our study.¹⁰

This is the eighth paper in a series of articles by the authors researching the differences between financial statement income (often called book income) and tax income as reported on U.S. corporate income tax returns.¹¹ This eighth paper compares 2011–2012 Schedule M-3 and Form 1120 tax return data profiles for Schedule

⁶ A major problem with interpreting Schedule M-1 data in the past was the fact that the taxpayer was allowed to report a starting Schedule M-1, line 1, book income amount without reconciling the reported book income amount to financial accounting income on the taxpayer's financial statements. *Schedule M-3, Part I, line 11, defines the starting book income for the book-to-tax reconciliation in Parts II and III.* The May 10, 2013, IRS notice, effective December 31, 2014, permitting the use of Schedule M-1 by corporations and partnerships with \$10 million but less than \$50 million in assets in place of Schedule M-3, Parts II and III, requires Schedule M-3, Part I, and requires that Schedule M-1, line 1, book income must equal Schedule M-3, line 11.

⁷ We define “SEC 10K/Public” to include any tax return on which: (1) Schedule M-3, Part I, line 1a, indicated that an SEC 10K was filed; or (2) Part I, line 3a, indicated that the corporation had publicly traded common stock. Some firms indicate the first without the second which may mean publicly traded debt or a reporting error. Other firms report the second without the first suggesting a reporting error. We make use of the presence of either indicator. We define “Audited” to include any tax return on which Schedule M-3, Part I, line 1b, indicates that certified audited financial statements were prepared and our requirements for “SEC 10K/Public” are not met. We define “Unaudited” to include all other returns.

⁸ Temporary differences are important in tax administration because they may identify that an item is being included in the wrong tax year. For example, deferring the recognition of \$1 billion of income for 30 years (or accelerating the recognition of \$1 billion of deductions by 30 years) involves a substantial time value of money change in the value of the tax due. In contrast to temporary differences, permanent differences are adjustments that arise as a result of fundamental permanent differences in financial and tax accounting rules. These differences result from transactions that will not reverse in subsequent periods. In financial statements reporting under generally accepted accounting principles (GAAP), permanent differences are not considered in the FAS No.109 (ASC 740) computation of deferred tax assets and liabilities, but do have a direct impact on the effective tax rate. Therefore, permanent differences have the potential to substantially influence reported financial earnings per share computations, and, in the case of public companies, stock prices. Accordingly, permanent differences of a given size may represent a greater audit risk than temporary differences of the same size.

⁹ See Section 3 of this report for a discussion of sign conventions.

¹⁰ Some companies with assets less than \$10 million voluntarily filed Schedule M-3. We do not analyze that data. Our minimum reconciliation tests require Schedule M-3 data agreement within tolerances of 1 percent of the maximum absolute value of the amounts on Part II, line 30, for income between Part I, line 11, and Part II, line 30, column (a), and for expenses/deductions between Part III, line 38 (line 36 through 2009), and the carryover line Part II, line 27. *In addition, effective for data from the 2012 SOI corporate file, we require that the amounts reported on Part I, lines 4a through line 10, reconcile with the total amount reported on line 11. If Part I, lines 4a to line 10, are blank, we set Part I, line 4a, worldwide income, to line 11, book income.* The reconciliations of the subset of corporations meeting our minimum data and reconciliation tests for this 2011–2012 Schedule M-3 study with the full 2011 and 2012 SOI corporate files are presented in Distribution Table D3 of the full M-3 First Look data sets for 2011 and 2012 available on request.

¹¹ See Boynton, DeFilippes, and Legel (2005, 2006a, 2006b, and 2008), Boynton, DeFilippes, Legel, and Reum (2011 and 2014) and Boynton, DeFilippes, Legel, and Rupert (2014). The first two articles analyze corporate Form 1120, Schedule M-1, reporting for Tax Years 1990–2003. The third paper in this series analyzes advance file data for the 2004 corporate Form 1120, Schedule M-3. The fourth paper analyzes final data for the 2005 corporate Form 1120, Schedule M-3, and updates the prior 2004 report using final 2004 data. The fifth paper analyzes final data for the 2006 and 2007 corporate Form 1120, Schedule M-3, as well as earlier Schedule M-1 data from 1994 through 2005 and Schedule M-3 data from 2004–2005. The sixth paper analyzes final data for 2008, 2009, and 2010 corporate Form 1120, Schedule M-3, as well as earlier Schedule M-3 data for 2006 and 2007 and information on 2010 Schedule UTP filing status. The seventh paper analyzes the Schedule M-3 profiles of Schedule UTP filers and nonfilers with \$100 million or more in assets in 2010 and 2011.

UTP (Uncertain Tax Position Statement) filers and nonfilers.¹² For Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets, the study further compares Schedule M-3 profiles of corporations that cite or do not cite, on Schedule UTP, Part I, any of the five most commonly cited Internal Revenue Code (IRC) sections: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized cost).

2. Schedule M-3 versus Schedule UTP

Schedule M-3 was introduced in 2004 for corporations with assets of \$10 million or more in order to assist the IRS in reconciling financial statement income to tax income including identifying temporary and permanent BTD. Taxpayers prepare corporate and partnership tax returns by adjusting amounts from their financial statements. The goal is to increase taxpayer transparency with respect to the adjustments made to the financial statements to prepare the tax return. Many but not all of the items that must be listed on Schedule UTP generate or impact BTD that must be included on Schedule M-3. Schedule M-3 reports dollar amounts; Schedule UTP does not.

Schedule UTP was introduced in 2010 for corporations with assets of \$100 million or more with audited financial statements reporting uncertain tax positions in the income tax footnote and for certain related corporations.¹³ The purpose was to share with the IRS some of the taxpayer information calculated as part of preparing the income tax footnote for the financial statements.¹⁴ The goal was to increase taxpayer transparency with respect to items giving rise to Federal income tax UTPs in the taxpayer's financial statements.

Schedule UTP asks for relevant code sections and a concise description of issues, *without dollar amounts*, for the UTPs that affect the financial statements' reported U.S. Federal income tax liabilities of certain corporations that issue or are included in audited financial statements. The corporate asset reporting threshold is assets of \$100 million or more in tax years beginning in 2010 and 2011, \$50 million or more in tax years beginning in 2012 and 2013, and \$10 million or more in tax years beginning in 2014 or later.¹⁵

Items listed on the Schedule UTP relate to amounts and/or positions reported on other forms or schedules of the current tax return or a prior tax return. Many of the Schedule UTP items relate to the temporary or permanent BTD reported on Parts II and III of the Schedule M-3. (Note that adjustments can be made during an examination for amounts reported on Part I of the Schedule M-3 due to errors in the calculations of the income/(loss) of the includible and excludable entities. However, it is unlikely taxpayers would report a Schedule UTP item that would relate to whether an entity should be included or excluded from the U.S. consolidated tax return group.) Other Schedule UTP items may relate to tax credit amounts or international issues that are not reported on the Schedule M-3, but are instead reported on the forms and schedules specific to those items (i.e., Form 6765, *Credit for Increasing Research Activities*, or Form 5471, *Information Return of U.S. Persons with Respect to Certain Foreign Corporations*).¹⁶

¹² The asset threshold that requires filing of Schedule UTP is \$100 million or more, effective for tax years beginning in 2010 or later, \$50 million or more for tax years beginning in 2012 or later, and \$10 million or more for tax years beginning in 2014 or later. Other conditions also apply as discussed in the next section. Except as otherwise indicated, this study is limited to Schedule UTP filers and nonfilers with SEC 10K/Public financial statements and \$100 million or more in assets. A Schedule UTP filer may file because it is required to file or because it files voluntarily. A Schedule UTP nonfiler may be a nonfiler either because it is not required to file or because it is required to file but failed to file. For tax years beginning in 2012 and later, the asset threshold for Schedule UTP drops to \$50 million or more. This study does not include the new asset range of \$50 million to \$100 million for Schedule UTP in the detailed analysis in Part III because a 2-year comparison is not possible.

¹³ For a discussion of the uncertain tax positions reported on Schedule UTP and an analysis of how Schedule UTP reporting requirements affect corporate tax and financial reporting behavior, see Towery (2015).

¹⁴ Footnote reporting of uncertain tax positions is required by U.S. GAAP under FAS No. 109 (ASC 740) and FIN 48 (ASC 740–10).

¹⁵ Schedule UTP requires the reporting of each U.S. Federal income tax position taken by an applicable corporation on its U.S. Federal income tax return for which two conditions are satisfied: (1) The corporation has taken a tax position on its U.S. Federal income tax return for the current tax year or for a prior tax year; and (2) either the corporation or a related party has recorded a reserve with respect to that tax position for U.S. Federal income tax in audited financial statements, or the corporation or related party did not record a reserve for that tax position because the corporation expects to litigate the position. A tax position for which a reserve was recorded (or for which no reserve was recorded because of an expectation to litigate) must be reported regardless of whether the audited financial statements are prepared based on U.S. GAAP, International Financial Reporting Standards (IFRS), or other country-specific accounting standards, including a modified version of any of the above (for example, modified GAAP).

¹⁶ Although Schedule M-3 does not deal with credits, a direct correlation may exist between an item on Schedule M-3 and a credit. For example, there is a direct correlation between Schedule M-3 R&D costs on Part III, line 36, column (d), and credit-eligible expenses on Form 6765, and therefore with the R&D credit. IRC section 41 relating to the R&D credit is the most frequent code section cited in 2011 and 2012 on Schedule UTP, Part I. The second frequent code section cited is IRC section 482 relating to transfer pricing.

Summary

- Schedule M-3:
 - Schedule M-3 is a crosswalk from the taxpayer's financial statements to their tax return.
 - Part I removes the income (loss) of all entities included in the financial statements but not included in the consolidated tax return and adds the income (loss) of all entities not included in the financial statements but included in the consolidated tax return.
 - Parts II and III require taxpayers to report the dollar amounts of the temporary and permanent adjustments they make to create their tax return from their financial statements as well as the initial book income and final tax income amounts for each scheduled item.
- Schedule UTP:
 - Schedule UTP reports the Federal income tax UTPs reserved on the taxpayer's financial statements with respect to items on the tax return the taxpayer acknowledges the IRS may challenge.
 - Schedule UTP discloses relevant IR Code sections and provides a concise description of the UTPs *without reporting the dollar amounts*.
 - Items listed on Schedule UTP may relate to the amounts reported on Schedule M-3.
 - Some items reported on Schedule UTP may relate to items not reported on Schedule M-3 (i.e., tax credit items).¹⁷

Schedule M-3 and Schedule UTP are complementary sources of taxpayer transparency that do not overlap and do not contain duplicative information.

3. Book-to-Tax Differences (BTD) and Signs

Book income is the financial statement income of the entity filing a corporation or partnership income tax return. For consolidated corporations filing U.S. Form 1120, book income is the consolidated financial statement income of the includible corporations joining in the consolidated tax return and will often differ from the worldwide consolidated income reported by the parent corporation's worldwide consolidated financial statements. Schedule M-3, Part I, reconciles worldwide consolidated financial statement income to book income.

We compare pretax book income (book income measured before Federal income tax expense) with tax income and calculate BTD as pretax differences, consistent with the BTD literature since Talisman (2000).¹⁸

The BTD literature prior to the introduction of Schedule M-3 defined the sign of the difference between pretax book income and tax income as “book minus tax” resulting in a positive difference if the book amount is *higher* than the tax amount. Schedule M-3 reverses this prior convention to “tax minus book” by its reconciliation rules.

For Schedule M-3 the temporary and permanent adjustment amounts reported in columns (b) and (c) of Parts II and III are the amounts that are *added* to column (a) book income to determine column (d) tax income. A *positive* total BTD in columns (b) and (c) of Schedule M-3, Parts II and III, means that the tax amount is *higher* than the book amount. A *negative* total BTD in columns (b) and (c) of Schedule M-3, Parts II and III, means that the tax amount is *lower* than the book amount.

In our report the sign of Schedule M-3, Part III, expense/deduction data including BTD has been changed to agree with the effect of such expense/deduction items and BTD on net income reported on Part II, line 30. If a Part III expense/deduction item or BTD reduces Part II, line 30, net income, it is shown as a negative amount in our report.¹⁹

¹⁷ See prior footnote on R&D credits and Schedule M-3.

¹⁸ We calculate total pretax book income and total pretax temporary and permanent BTD by adding back Federal income tax expense and differences reported on Schedule M-3, Part III, lines 1 and 2, columns (a), (b), and (c), to book income and differences reported on Schedule M-3, Part II, line 30, columns (a), (b), and (c), column by column. Total BTD is the sum of total temporary and permanent BTD.

¹⁹ Schedule M-3 instructions require that column (a) book expense and column (d) tax deduction amounts that reduce net book income and reduce net tax income be shown on Part III as positive amounts. However, some taxpayers fail to follow the instructions. For a discussion of the problem and how we deal with it, see Boynton, DeFilippes, and Legel (2006b and 2008) and Boynton, DeFilippes, Legel, and Reum (2011).

4. Source of Schedules M-3 Data and UTP Status

A weighted statistical sample of tax return data is electronically encoded annually by the IRS Statistics of Income (SOI) Division for use by Treasury's Office of Tax Analysis (OTA), and the congressional Joint Committee on Taxation (JCT).²⁰ Planning, Analysis, Inventory, and Research (PAIR), within the IRS Large Business & International Division (LB&I), also receives a copy of the file.²¹ The SOI corporate file includes Schedule M-1 data and, beginning with the 2004 file, Schedule M-3 data. Starting with 2010, the SOI corporate file reports if the taxpayer indicates on Form 1120, Schedule K, that Schedule UTP is required, if a Schedule UTP, Part I, identifying a UTP is attached to the return with any data, and the number of lines on Schedule UTP, Part I, with any data on the line.²² The 2012 SOI corporate file was issued to OTA, JCT, and LB&I in October 2014.²³

Beginning May 2011, researchers using SOI data must report tax data as an aggregate for a minimum of five taxpayers to protect taxpayer confidentiality.²⁴ For statistical reasons, SOI prefers that reported aggregate data are for 10 or more taxpayers when possible.²⁵

5. Limits of Schedule M-3 Data

With the exception of Schedule M-3, Part I, amounts reported on the Form 1120 tax return and the Schedule M-3, Parts II and III:

- are limited to the tax information and pretax book income information of the includible corporations in the tax consolidated return; and
- do not include the tax information or pretax book income information of the nonincludible corporations and partnerships (both foreign and domestic) that are included in the worldwide consolidated after-tax income reported on Schedule M-3, Part I, line 4 (the worldwide book income reported in the financial statements for consolidated book purposes).

The after-tax income of the nonincludible corporations and partnerships is removed, in gross after-tax amounts, on Schedule M-3, Part I, lines 5 and 6, as one step in determining the book income of the includible corporations reported on Schedule M-3, Part I, line 11.

Form 1120 tax return and Schedule M-3 data do not yield generalizations about the financial statement pretax consolidated worldwide income. In particular, amounts reported on Form 1120 and Schedule M-3 do not provide the data needed to calculate the pretax worldwide effective tax rate for the entities included in the worldwide financial statements.

²⁰ The SOI corporate file is a statistical sample. The record for a smaller tax return (usually measured by total assets) may be weighted to represent more than one tax return. Generally tax returns for corporations with \$50 million or more in assets have a weight of one, that is, the record represents only itself. The record for a smaller tax return generally has a weight greater than one (for example five), that is, the record represents several similar tax returns (for example, five tax returns). The SOI corporate data file for year t includes all tax years ending between July of Calendar Year t and June of Calendar Year $t+1$.

²¹ Use of the SOI file by PAIR and LB&I is limited under a formal Memorandum of Understanding between SOI and LB&I to research studies. SOI file data are not used for IRS audit case building.

²² The regular 2011 and 2012 SOI corporate files do not tabulate what is reported on Schedule UTP, Part I, and do not report if an attached Schedule UTP, Part I, contains relevant data. A special SOI supplement to each of the regular 2011 and 2012 SOI corporate files tabulates the limited information reported on Schedule UTP, Part I, lines 1 through 10, for current year UTPs such as IRC sections cited, temporary and permanent effect, whether the position is a major position, and relative rank of the position. Part II, relating to prior-year UTPs, and Part III, relating to the concise descriptions for the positions listed in Parts I and II, are not tabulated by SOI.

²³ The final SOI corporate file may contain placeholder records representing returns for some reason not available at the time the SOI file is issued but desired by SOI for statistical purposes. Placeholder data are commonly the edited return data from the prior tax year, but may also be current-year data from the IRS Business Master File (limited return data tabulated by the IRS when the return is first received and processed) or data from the IRS Employee User Portal. Placeholder returns are not included in the Schedule M-3 First Look data files.

²⁴ Prior to May 2011, the minimum aggregation requirement for SOI and for other government agencies was data aggregation for three or more taxpayers or individuals. SOI has increased the required minimum for the use of SOI data to five or more. The change for SOI data applies to Tax Year 2008 and to new studies of data from earlier tax years. A data count of zero is permitted. Tests must be performed to assure that data cannot be generated by subtraction that would violate the minimum aggregation requirement. For a discussion of the older requirement of three or more taxpayers or individuals for aggregate data, see U.S. Office of Management and Budget Working Paper 22 (2005) and IRS Publication 1075 (Rev 2007).

²⁵ Our tax return table values may not add and may differ from official 2011 and 2012 SOI values due to rounding. SOI publications do not include Schedules M-1 or M-3 data. Prior to the publication of Boynton, DeFilippes, and Legel (2005 and 2006a), only Plesko (2002) (for 1996–1998) and Plesko and Shumofsky (2005) (for 1995–2001) presented public Schedule M-1 data for the SOI corporate file population. The year-by-year reconciliations of the subset of corporations meeting our minimum data and reconciliation tests for this 2011–2012 Schedule M-3 study with the full 2011–2012 SOI corporate files are presented in Distribution Table D3 of the full M-3 First Look data set for each year, 2011–2012, available on request. Our minimum data and reconciliation tests require that Part I, line 11, and Part II, line 30, column (a), agree and that Part III, line 38, and Part II, line 27, agree within 1 percent of the maximum absolute value of the amounts on Part II, line 30. In addition, effective for data from the 2012 SOI corporate file, we require that the amounts reported on Part I, line 4a through line 10, reconcile with the total amount reported on line 11. If Part I, line 4a to line 10, are blank, we set Part I, line 4a, worldwide income to line 11, book income.

6. Reconciling Counts of Schedule UTP

The number of Schedule UTP reported in this paper for Tax Years 2011 and 2012 differ from the number reported by the LB&I Schedule UTP Web page for the 2011 and 2012 form years.²⁶ The differences are a result of: (1) including different corporate income tax return forms (LB&I counts include Form 1120-F and Forms 1120, 1120-L, and 1120-PC filed by parents of insurance companies, and this report does not); (2) using different tax year ending months for Tax Years 2011 and 2012 (for example, July 2012 to June 2013 for SOI Tax Year 2012 versus December 2012 to November 2013 for LB&I Form Year 2012); (3) using different standards as to whether Schedule UTP is filed (this study requires only that either SOI reports that on Form 1120, Schedule K, line 10, the taxpayer indicates Schedule UTP is required or SOI indicates Schedule UTP, Part I, Current Year, is present and has one or more lines with data; LB&I requires both Schedule UTP, Parts I, Current Year, or II, Prior Year, and also requires Part III, Concise Descriptions);²⁷ and (4) different minimum asset recognition thresholds (LB&I includes Schedule UTP filing by a corporation without regard to assets and, except as otherwise noted, this study includes only corporations with assets of \$100 million or more).

The frequency of Schedule UTP filers citing an IRC section reported in this paper for Tax Years 2011 and 2012 differ from the frequency reported by the LB&I Schedule UTP Web page for the 2011 and 2012 form years for the frequency of UTPs described by an IRC section. The differences are a result of: (1) differences discussed in the prior paragraph in the Schedule UTP included for Tax Years 2011 and 2012 for this study and the Schedule UTP included in Form Years 2011 and 2012 for the LB&I Schedule UTP Web page; (2) this study reports the frequency with which an IRC section is cited by any Schedule UTP filer for any UTP (note that taxpayers may cite up to three IRC sections for any UTP and may cite a particular IRC section for more than one UTP); and (3) the LB&I Schedule UTP Web page reports the frequency with which UTP concise statements across all Schedule UTP filers are categorized in terms of a single most descriptive IRC section. (The LB&I Schedule UTP team determines a single most descriptive IRC section for each UTP concise description, which may differ from the IRC sections cited by the taxpayer.)

7. Mini M-3: Specified versus Other Lines

The “other-with-difference” lines on Schedule M-3 with BTD are Part II, line 25, and Part III, line 37. The “other-with-no-difference” line is Part II, line 28. In two prior studies in this series we noted both the large dollar magnitude of the book income, tax income, and BTD amounts reported on the “other-with-difference” lines and the documentation problems found on the lines.²⁸

We use a “Mini M-3” format to compare the aggregate amounts reported on the Schedule M-3, Parts II and III, “other-with-difference” or “other-with-no-difference” lines, with the aggregate amounts reported on the Schedule M-3, Parts II and III, “specified” lines, that is, the lines with specific captions.²⁹

A Schedule M-3 cost-of-goods-sold (COGS) adjustment discussed in the next section is used to remove the cost of securities, commodity contracts, and other financial products reported in COGS by some corporations and to reconcile to the COGS amount reported by the IRS SOI corporate data file. The Mini M-3 format also makes related special adjustments to other-income-with-difference and other-items-with-no-difference lines and decomposes the adjusted other-items-with-no-difference line into other-income-with-no-difference and other-expense/deduction-with-no-difference lines. Finally, the other-income-with-no-difference line is adjusted to reconcile to the gross receipts amount reported by the IRS SOI corporate data file.

²⁶ The LB&I Schedule UTP Web page is at <http://www.irs.gov/Businesses/Corporations/UTPFilingStatistics>.

²⁷ The regular 2011 and 2012 SOI corporate files do not tabulate what is reported on Schedule UTP, Part I, and do not report if an attached Schedule UTP, Part I, contains relevant data. A special SOI supplement to each of the regular 2011 and 2012 SOI corporate files tabulates the limited information reported on Schedule UTP, Part I, lines 1 through 10, for current year UTPs, such as IRC sections cited, temporary and permanent effect, whether the position is a major position, and relative rank of the position. Part II, relating to prior-year UTPs, and Part III, relating to the concise descriptions for the positions listed in Parts I and II, are not tabulated by SOI.

²⁸ For discussions of the other-with-difference documentation by large taxpayers in 2005 and 2007, see Boynton, DeFilippes, and Legel (2008) and Boynton, DeFilippes, Legel, and Reum (2011).

²⁹ Amounts reported on the other-with-difference lines require attached documentation. The documentation must separately state and adequately disclose the BTD for the line. The other-items-with-no-difference line has no documentation. Reporting on the other-with-difference lines is similar to but more detailed than reporting on Schedule M-1. Both allow descriptions determined by the taxpayer. Schedule M-1 requires only a description and a BTD. Schedule M-3 requires a description, a book income amount, a temporary BTD amount, a permanent BTD amount, and a tax income amount.

After making the data adjustments, the Mini M-3 format has 10 categories of specified lines, other-with-difference or no-difference lines, and subtotals or totals:³⁰

- Other income with no difference (Part II, line 28, adjusted) (gross receipts);
- COGS (Part II, line 17, adjusted);
- Adjusted gross profit;
- Specified income (Part II, lines 1–16, 18–24, and 29a–29c);
- Other income with difference (Part II, line 25, adjusted);
- Adjusted total income;
- Specified expense/deduction (Part III, lines 3–36);³¹
- Other expense/deduction with difference (Part III, line 37);
- Other expense/deduction with no difference (an adjustment to Part II, line 28); and
- Pretax book income.

We used the adjusted-total-income book amount as a common-size scaling factor and compare percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers and nonfilers for the three financial statement types, total pretax income BTD is expressed as a percentage of total pretax book income.

8. COGS and Other Adjustments

We make a Schedule M-3 COGS adjustment for the Mini M-3. The adjustment reconciles the Schedule M-3 COGS tax income amount with Form 1120, page 1, line 2, COGS reported by SOI for the corporations in our study. SOI removes the cost of securities, commodity contracts, and other financial products reported in Form 1120, page 1, line 2, COGS.³² We make the equal adjustments to Part II, line 17, COGS book income and tax income with the result that COGS BTD are not changed. SOI also makes adjustments to Form 1120, page 1, line 1, gross receipts to match the amounts SOI removes from COGS. We match our COGS adjustments with adjustments to other-income-with-difference and to other-items-with-no-difference. We also decompose the adjusted other-items-with-no-difference into other-income-with-no-difference and other-expense/deduction-with-no-difference.³³ Finally, the other-income-with-no-difference line is adjusted to reconcile to the gross receipts amount reported by the SOI corporate data file.

SOI has adjusted Form 1120, page 1, line 1, gross receipts and line 2, COGS; Schedule A, COGS; and Schedule L, inventory amounts, since the 1980s to remove the cost of securities and commodities transactions. SOI-adjusted COGS, gross receipts, and inventory amounts are used by the Bureau of Economic Analysis (BEA) for national income accounts. At the request of OTA, SOI has not adjusted Schedule M-3 data since their introduction in 2004.

We wish to develop a consistent Schedule M-3 measure of total book income before expenses to scale or common-size book income and tax income components and book expense and tax deduction components for different size corporations. Adopting the SOI adjustments to COGS and gross receipts facilitates development of a consistent measure of total income applicable to different size corporations.³⁴

³⁰ All BTD in adjusted gross profit are from adjusted COGS. The adjustments we make to COGS are made equally to the unadjusted book amount and tax amount and have no effect on the BTD.

³¹ We exclude Federal income tax expense reported on Schedule M-3, Part III, lines 1 and 2, from our pretax analysis. See our discussion of pretax income and BTD in Section 3 of this study.

³² Note that changes on the SOI corporate file do not change the amounts on the tax return and do not impact IRS audits (or lack of audits) for corporate tax returns.

³³ We have introduced adjustment lines into our 2011–2012 M-3 First Look FORM tables to show the frequency of adjustment and the amounts needed to reconcile Schedule M-3, Part II, line 17, COGS to the SOI amount reported for Form 1120, page 1, line 2.

³⁴ Aggregate unadjusted book income and tax income reported on Schedule M-3, Part II, line 26, for all corporations are both *negative* because the large absolute amount of COGS for all corporations on Part II, line 17, exceeds the income reported on the specified income lines and the other-income-with-difference line combined. A majority of gross receipts is reported on Part II, line 28, other items, with no difference.

We adjust 2012 Schedule M-3 COGS book and tax income amounts to agree with the SOI Form 1120, page 1, line 2, COGS. In doing so, we need to determine where on Schedule M-3 to make the matching gross receipts adjustment. Using 2010 data, we developed a rule to allocate the matching gross receipts reduction between Schedule M-3, Part II, line 25, other income with difference, and line 28, other items without difference.³⁵

In addition, we compare the Form 1120, page 1, line 27, total deduction amount with the total Part III deduction amount carried over to Part II as reported on Part II, line 27, column (d), to determine the total deductions-with-no-difference amount currently included in Part II, line 28, other items with no difference.³⁶ We then decompose the adjusted other-items-with-no-difference book and tax income amounts into other-income-with-no difference and other-expense/deduction-with-no difference. Finally, the other-income-with-no difference book and tax income amounts are adjusted to reconcile to the gross receipts amount reported by the SOI corporate data file on Form 1120, page 1, line 1, with a matching adjustment to the other-income-with-difference book and tax income amounts on Schedule M-3, Part II, line 25.

The adjustments do not affect pretax net income and do not affect BTD. BTD are not affected by the COGS and other adjustments described because equal adjustments are made to book income and tax income amounts.

We will use the adjusted book income and tax income amounts in our Mini M-3 analysis in Part III of this study and will scale by adjusted total income, the sum of the adjusted other-income-with-no-difference, adjusted COGS, specified-income, and adjusted other-income-with-difference amounts.

Part II. 2011–2012 U.S. Corporation Overview: Asset Size, FS Type, and UTP Status

Part II of this study describes the general population of Schedule UTP filers and nonfilers in terms of asset ranges (\$10 million to \$100 million, \$100 million to \$1 billion, and \$1 billion or more), by financial statement type (SEC 10K/Public, Audited, Unaudited), and by Schedule UTP filing status (filer or nonfiler).³⁷

Part III of this study will focus on data describing characteristics of Schedule UTP filers and nonfilers having SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more, but with Audited (but not SEC 10K/Public) or Unaudited financial statements.³⁸

The LB&I Division of the IRS is responsible for auditing corporations and partnerships with \$10 million or more in assets. Tables 1A and 1B along with Tables 2A and 2B highlight relevant Form 1120 corporation income tax return data for LB&I taxpayers with assets greater than \$10 million for years 2011 and 2012, respectively.

³⁵ We verified our rule on the 2010 data using the top 25 returns which, for 2010, accounted for 99 percent of the aggregate adjustment of approximately \$32 trillion. See Boynton, DeFilippes, Legel, and Reum (2014).

³⁶ Our allocation rule:

ADJ COGS1 and ADJ COGS2: If the absolute value of P2L17 column D COGS is greater than Form 1120, page 1, line 2, COGS, then the excess difference is the COGS adjustment and the matching gross receipts adjustment. The adjustments reduce the absolute magnitude of P2L17, P2L25, and P2L28.

ADJ COGS1: The gross receipts adjustment is applied to P2L25 other income with difference if P2L25D other income with difference is greater than P2L28D other income without difference AND P2L25D is greater than 80 percent of the gross receipts adjustment

ELSE use

ADJ COGS2: The gross receipts adjustment goes to P2L28 other income without difference.

ADJ COGS3: If the absolute value of [P2L17 column (d) COGS] is less than [1120, page 1, line 2, COGS], the adjustment is an increase to P2L17 and P2L28 in absolute magnitude.

ADJ EXPDED: We estimate expenses/deductions without differences as the amount if any by which Form 1120, page 1, line 27, total deductions exceed the absolute value of P2L27 column (d). We show it as an additional expense/deduction line and as an increase to P2L28. The adjusted P2L28 amount changes from “other items without difference” to “other income without difference.”

³⁷ A Schedule UTP filer may file because it is required to file or because it files voluntarily. A Schedule UTP nonfiler may be a nonfiler either because it is not required to file or because it is required to file but failed to file.

³⁸ For tax years beginning in 2012 and later, the asset threshold for Schedule UTP drops to \$50 million or more. This study does not include the new asset range of \$50 million to \$100 million for Schedule UTP in the detailed analysis in Part III because a 2-year comparison is not possible.

TABLE 1A. UTP Presence by Asset Size and Financial Statement for Tax Year 2011 (Dollar amounts in millions)

	Returns		Total Assets		Worldwide Income (Part I, Line 4)		Nonincludible Foreign Income		Pretax Book		Tax Net Income		Tax After Credits		Foreign Tax Credit	
	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%
All	41,636	100%	51,095,020	100%	820,641	100%	-776,435	100%	879,279	100%	715,873	100%	190,756	100%	104,882	100%
No Schedule UTP Filed	39,476	95%	15,207,307	30%	50,936	6%	-80,134	10%	76,819	9%	104,741	15%	15,946	34%	15,946	15%
Schedule UTP Filers	2,160	5%	35,887,713	70%	769,705	94%	-696,301	90%	802,460	91%	611,132	85%	125,278	66%	88,936	85%
Assets > \$100M	12,307	30%	50,193,758	98%	821,100	100%	-768,079	99%	882,033	100%	721,723	101%	180,792	95%	104,489	100%
No Schedule UTP Filed	10,233	25%	14,310,291	28%	51,096	6%	-72,253	9%	78,817	9%	110,125	15%	55,568	29%	15,577	15%
Schedule UTP Filers	2,074	5%	35,883,467	70%	770,004	94%	-695,826	90%	803,216	91%	611,598	85%	125,224	66%	88,912	85%
IRC 41 Cited	894	2%	12,968,969	25%	401,960	49%	-447,672	58%	396,657	45%	282,142	39%	60,800	32%	27,412	26%
IRC 482 Cited	580	1%	11,900,515	23%	436,914	53%	-547,900	71%	362,366	41%	346,609	48%	51,918	27%	61,459	59%
IRC 162 Cited	266	1%	6,835,859	13%	236,615	29%	-107,329	14%	266,196	30%	202,600	28%	35,093	18%	38,649	37%
IRC 199 Cited	184	0%	3,555,282	7%	219,551	27%	-164,848	21%	222,904	25%	209,271	29%	31,877	17%	38,495	37%
IRC 263 Cited	121	0%	3,203,219	6%	105,703	13%	-74,662	10%	131,983	15%	87,612	12%	18,169	10%	16,498	16%
Assets > \$1B	2,751	7%	47,199,385	92%	782,195	95%	-753,264	97%	833,435	95%	687,664	96%	157,550	83%	101,499	97%
No Schedule UTP Filed	1,679	4%	11,711,321	23%	24,801	3%	-64,184	8%	41,953	5%	88,008	12%	37,266	20%	13,533	13%
Schedule UTP Filers	1,072	3%	35,488,064	69%	757,395	92%	-689,080	89%	791,482	90%	599,656	84%	120,283	63%	87,966	84%
IRC 41 Cited	412	1%	12,775,960	25%	395,007	48%	-444,033	57%	390,541	44%	276,399	39%	58,363	31%	27,010	26%
IRC 482 Cited	360	1%	11,806,232	23%	432,743	53%	-544,478	70%	358,381	41%	342,434	48%	50,789	27%	60,883	58%
IRC 162 Cited	203	0%	6,805,328	13%	235,094	29%	-106,177	14%	264,557	30%	200,591	28%	34,488	18%	38,502	37%
IRC 199 Cited	123	0%	3,528,857	7%	216,904	26%	-163,805	21%	220,023	25%	206,980	29%	31,248	16%	38,362	37%
IRC 263 Cited	100	0%	3,193,308	6%	105,202	13%	-74,655	10%	131,367	15%	87,342	12%	18,079	9%	16,465	16%
\$100M < Assets < \$1B	9,556	23%	2,994,373	6%	38,905	5%	-14,814	2%	48,597	6%	34,058	5%	23,241	12%	2,990	3%
No Schedule UTP Filed	8,554	21%	2,598,970	5%	26,295	3%	-8,068	1%	36,863	4%	22,116	3%	18,301	10%	2,044	2%
Schedule UTP Filers	1,002	2%	395,403	1%	12,610	2%	-6,746	1%	11,734	1%	11,942	2%	4,940	3%	946	1%
IRC 41 Cited	482	1%	193,009	0%	6,942	1%	-3,639	0%	6,116	1%	5,743	1%	2,437	1%	402	0%
IRC 482 Cited	220	1%	94,283	0%	4,171	1%	-3,422	0%	3,985	0%	4,175	1%	1,130	1%	576	1%
IRC 162 Cited	63	0%	30,531	0%	1,521	0%	-1,152	0%	1,639	0%	2,009	0%	605	0%	147	0%
IRC 199 Cited	61	0%	26,425	0%	2,647	0%	-1,043	0%	2,881	0%	2,291	0%	628	0%	133	0%
IRC 263 Cited	21	0%	9,911	0%	501	0%	-7	0%	616	0%	270	0%	91	0%	33	0%
\$10M < Assets < \$100M	29,329	70%	901,262	2%	-459	0%	-8,357	1%	-2,754	0%	-5,850	-1%	9,965	5%	393	0%
No Schedule UTP Filed	29,243	70%	897,016	2%	-160	0%	-7,881	1%	-1,998	0%	-5,384	-1%	9,911	5%	369	0%
Schedule UTP Filers	86	0%	4,246	0%	-299	0%	-475	0%	-756	0%	-467	0%	54	0%	24	0%
SEC 10K/Public																
All	4,488	11%	7,697,048	74%	867,248	106%	-748,596	96%	878,771	100%	637,691	89%	137,090	72%	94,412	90%
No Schedule UTP Filed	3,250	8%	6,346,372	12%	134,769	16%	-65,676	8%	138,573	16%	79,351	11%	28,086	15%	11,170	11%
Schedule UTP Filers	1,238	3%	31,350,676	61%	732,479	89%	-682,920	88%	740,198	84%	558,340	78%	109,004	57%	83,242	79%
Assets > \$100M	1,227	3%	31,349,918	61%	732,811	89%	-682,842	88%	740,593	84%	558,750	78%	108,995	57%	83,242	79%
IRC 41 Cited	594	1%	12,017,709	24%	382,692	47%	-445,473	57%	361,667	41%	256,933	36%	54,004	28%	26,484	25%
IRC 482 Cited	335	1%	9,422,154	18%	418,507	51%	-542,484	70%	332,959	38%	319,426	45%	44,339	23%	58,663	56%
IRC 162 Cited	188	0%	6,354,718	12%	235,892	29%	-106,759	14%	265,144	30%	200,297	28%	33,426	18%	38,577	37%
IRC 199 Cited	135	0%	3,238,242	6%	210,439	26%	-164,334	21%	203,127	23%	200,197	28%	29,594	16%	38,199	36%
IRC 263 Cited	93	0%	2,574,949	5%	106,345	13%	-74,654	10%	132,747	15%	84,418	12%	17,001	9%	16,292	16%
Assets > \$1B	748	2%	31,148,090	61%	723,340	88%	-677,128	87%	732,578	83%	551,856	77%	106,199	56%	82,680	79%
IRC 41 Cited	313	1%	11,901,987	23%	377,192	46%	-442,147	57%	357,238	41%	253,245	35%	52,472	28%	26,219	25%
IRC 482 Cited	246	1%	9,379,195	18%	415,226	51%	-539,710	70%	330,331	38%	317,175	44%	43,839	23%	58,332	56%
IRC 162 Cited	159	0%	6,339,808	12%	234,372	29%	-105,620	14%	263,577	30%	198,735	28%	32,963	17%	38,438	37%
IRC 199 Cited	96	0%	3,221,419	6%	208,264	25%	-163,359	21%	200,802	23%	198,261	28%	29,080	15%	38,101	36%
IRC 263 Cited	83	0%	2,569,772	5%	106,082	13%	-74,601	10%	132,482	15%	84,165	12%	16,937	9%	16,259	16%

TABLE 1B. UTP Presence by Asset Size and Financial Statement for Tax Year 2011 (Dollar amounts in millions)

	Returns		Total Assets		Worldwide Income (Part I, Line 4)		Nonincludible Foreign Income		Pretax Book		Tax Net Income		Tax After Credits		Foreign Tax Credit	
	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%
SEC-10K/Public Filers of Schedule UTP (continued)																
\$100M < Assets < \$1B	479	1%	201,828	0%	9,471	1%	-5,714	1%	8,015	1%	6,894	1%	2,797	1%	562	1%
IRC 41 Cited	281	1%	115,722	0%	5,500	1%	-3,326	0%	4,428	1%	3,688	1%	1,532	1%	265	0%
IRC 482 Cited	89	0%	42,959	0%	3,281	0%	-2,774	0%	2,628	0%	2,251	0%	500	0%	330	0%
IRC 162 Cited	29	0%	14,910	0%	1,520	0%	-1,139	0%	1,567	0%	1,562	0%	463	0%	139	0%
IRC 199 Cited	39	0%	16,823	0%	2,175	0%	-974	0%	2,325	0%	1,936	0%	504	0%	98	0%
IRC 263 Cited	10	0%	5,177	0%	263	0%	-53	0%	264	0%	252	0%	64	0%	33	0%
Assets < \$100M	11	0%	758	0%	-332	0%	-77	0%	-394	0%	-409	0%	8	0%	0	0%
Audited	17,298	42%	6,354,953	12%	63,688	8%	-22,616	3%	92,621	11%	45,882	6%	27,160	14%	5,769	6%
No Schedule UTP Filed	16,723	40%	3,791,167	7%	36,225	4%	-10,438	1%	48,223	5%	23,611	3%	20,233	11%	1,620	2%
Schedule UTP Filers	575	1%	2,563,786	5%	27,463	3%	-12,179	2%	44,398	5%	22,270	3%	6,927	4%	4,149	4%
Assets > \$100M	535	1%	2,561,995	5%	27,278	3%	-11,768	2%	44,613	5%	22,246	3%	6,914	4%	4,125	4%
Assets > \$1B	159	0%	2,428,108	5%	24,472	3%	-10,914	1%	41,776	5%	18,276	3%	5,364	3%	3,864	4%
\$100M < Assets < \$1B	376	1%	133,887	0%	2,805	0%	-853	0%	2,837	0%	3,970	1%	1,551	1%	261	0%
Assets < \$100M	39	0%	1,791	0%	185	0%	-411	0%	-215	0%	24	0%	13	0%	24	0%
Unaudited	19,850	48%	7,043,018	14%	-110,295	-13%	-5,223	1%	-92,114	-10%	32,300	5%	26,506	14%	4,701	4%
No Schedule UTP Filed	19,503	47%	5,069,768	10%	-120,058	-15%	-4,020	1%	-109,977	-13%	1,779	0%	17,159	9%	3,156	3%
Schedule UTP Filers	347	1%	1,973,250	4%	9,763	1%	-1,203	0%	17,864	2%	30,521	4%	9,347	5%	1,545	1%
Assets > \$100M	311	1%	1,971,553	4%	9,915	1%	-1,216	0%	18,011	2%	30,603	4%	9,314	5%	1,545	1%
Assets > \$1B	165	0%	1,911,866	4%	9,582	1%	-1,037	0%	17,128	2%	29,524	4%	8,721	5%	1,422	1%
\$100M < Assets < \$1B	146	0%	59,687	0%	333	0%	-178	0%	883	0%	1,079	0%	592	0%	123	0%
Assets < \$100M	36	0%	1,697	0%	-152	0%	13	0%	-147	0%	-82	0%	33	0%	0	0%
Non-Public Filers of Schedule UTP																
Assets > \$100M	846	2%	4,533,548	9%	37,193	5%	-12,983	2%	62,624	7%	52,849	7%	16,228	9%	5,670	5%
IRC 41 Cited	300	1%	951,260	2%	19,257	2%	-2,199	0%	34,990	4%	25,209	4%	6,796	4%	928	1%
IRC 482 Cited	245	1%	2,478,361	5%	18,407	2%	-5,416	1%	29,407	3%	27,182	4%	7,560	4%	2,796	3%
IRC 162 Cited	78	0%	481,141	1%	724	0%	-570	0%	1,052	0%	2,303	0%	1,667	1%	73	0%
IRC 199 Cited	49	0%	317,040	1%	9,112	1%	-515	0%	19,777	2%	9,074	1%	2,293	1%	296	0%
IRC 263 Cited	28	0%	628,270	1%	-642	0%	-8	0%	-764	0%	3,195	0%	1,168	1%	206	0%
Assets > \$1B	324	1%	4,339,974	8%	34,054	4%	-11,952	2%	58,904	7%	47,800	7%	14,085	7%	5,286	5%
IRC 41 Cited	99	0%	873,973	2%	17,815	2%	-1,886	0%	33,303	4%	23,155	3%	5,891	3%	791	1%
IRC 482 Cited	114	0%	2,427,037	5%	17,517	2%	-4,768	1%	28,050	3%	25,258	4%	6,950	4%	2,550	2%
IRC 162 Cited	44	0%	465,520	1%	722	0%	-556	0%	980	0%	1,857	0%	1,524	1%	64	0%
IRC 199 Cited	27	0%	307,438	1%	8,641	1%	-446	0%	19,222	2%	8,719	1%	2,168	1%	261	0%
IRC 263 Cited	17	0%	623,536	1%	-880	0%	-54	0%	-1,115	0%	3,177	0%	1,142	1%	206	0%
\$100M < Assets < \$1B	522	1%	193,575	0%	3,139	0%	-1,032	0%	3,720	0%	5,048	1%	2,143	1%	384	0%
IRC 41 Cited	201	0%	77,288	0%	1,442	0%	-313	0%	1,687	0%	2,055	0%	906	0%	137	0%
IRC 482 Cited	131	0%	51,324	0%	890	0%	-648	0%	1,357	0%	1,924	0%	629	0%	246	0%
IRC 162 Cited	34	0%	15,621	0%	1	0%	-13	0%	71	0%	447	0%	143	0%	8	0%
IRC 199 Cited	22	0%	9,603	0%	471	0%	-69	0%	556	0%	355	0%	124	0%	35	0%
IRC 263 Cited	11	0%	4,734	0%	238	0%	46	0%	351	0%	18	0%	27	0%	0	0%

TABLE 2A. UTP Presence by Asset Size and Financial Statement for Tax Year 2012 (Dollar amounts in millions)

	Returns		Total Assets		Worldwide Income (Part I, Line 4)		Nonincludible Foreign Income		Pretax Book		Tax Net Income		Tax After Credits		Foreign Tax Credit	
	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%
All	42,301	100%	50,897,122	100%	969,512	100%	-669,129	100%	1,027,879	100%	889,481	100%	228,822	100%	81,349	100%
No Schedule UTP Filed	40,069	95%	21,188,137	42%	277,831	29%	-87,779	13%	288,354	28%	235,288	26%	83,278	36%	16,285	20%
Schedule UTP Filers	2,232	5%	29,708,985	58%	691,680	71%	-581,350	87%	739,525	72%	654,193	74%	145,545	64%	65,064	80%
Assets > \$100M	12,427	29%	49,983,146	98%	963,762	99%	-660,668	99%	1,030,820	100%	886,343	100%	217,370	95%	80,912	99%
No Schedule UTP Filed	10,409	25%	20,288,995	40%	271,450	28%	-79,495	12%	290,672	28%	231,653	26%	72,184	32%	15,868	20%
Schedule UTP Filers	2,018	5%	29,694,251	58%	692,312	71%	-581,173	87%	740,148	72%	654,690	74%	145,186	63%	65,044	80%
IRC 41 Cited	772	2%	8,090,907	16%	276,498	29%	-247,853	37%	280,587	27%	218,371	25%	52,915	23%	14,168	17%
IRC 482 Cited	554	1%	9,469,654	19%	359,576	37%	-467,124	70%	292,600	28%	277,767	31%	55,008	24%	35,987	44%
IRC 162 Cited	247	1%	9,704,072	19%	164,886	17%	-66,778	10%	213,091	21%	167,669	19%	40,398	18%	15,480	19%
IRC 199 Cited	179	0%	2,417,199	5%	195,509	20%	-165,073	25%	152,274	15%	140,281	16%	30,823	13%	14,709	18%
IRC 263 Cited	128	0%	2,550,571	5%	95,932	10%	-63,867	10%	153,750	15%	92,352	10%	19,080	8%	13,348	16%
Assets > \$1B	2,856	7%	46,971,832	92%	924,509	95%	-648,604	97%	979,508	95%	830,805	93%	189,923	83%	78,418	96%
No Schedule UTP Filed	1,777	4%	17,661,062	35%	240,752	25%	-72,165	11%	250,620	24%	191,121	21%	50,237	22%	14,107	17%
Schedule UTP Filers	1,079	3%	29,310,570	58%	683,757	71%	-576,439	86%	728,888	71%	639,684	72%	139,666	61%	64,311	79%
IRC 41 Cited	336	1%	7,912,481	16%	271,784	28%	-245,475	37%	274,310	27%	210,782	24%	50,211	22%	13,859	17%
IRC 482 Cited	355	1%	9,378,297	18%	357,220	37%	-464,672	69%	290,308	28%	273,463	31%	53,768	23%	35,754	44%
IRC 162 Cited	190	0%	9,677,327	19%	164,318	17%	-66,397	10%	212,558	21%	166,824	19%	40,059	18%	15,451	19%
IRC 199 Cited	119	0%	2,389,447	5%	192,749	20%	-163,897	24%	149,121	15%	137,030	15%	29,921	13%	14,573	18%
IRC 263 Cited	99	0%	2,538,822	5%	95,806	10%	-63,811	10%	153,568	15%	92,098	10%	18,993	8%	13,313	16%
\$100M < Assets < \$1B	9,571	23%	3,011,515	6%	39,254	4%	-12,064	2%	51,313	5%	55,538	6%	27,446	12%	2,494	3%
No Schedule UTP Filed	8,632	20%	2,627,833	5%	30,699	3%	-7,330	1%	40,053	4%	40,531	5%	21,946	10%	1,761	2%
Schedule UTP Filers	939	2%	383,682	1%	8,555	1%	-4,734	1%	11,260	1%	15,006	2%	5,500	2%	733	1%
IRC 41 Cited	436	1%	178,425	0%	4,714	0%	-2,378	0%	6,277	1%	7,589	1%	2,703	1%	309	0%
IRC 482 Cited	199	0%	91,357	0%	2,356	0%	-2,453	0%	2,292	0%	4,304	0%	1,240	1%	243	0%
IRC 162 Cited	57	0%	26,745	0%	567	0%	-382	0%	533	0%	845	0%	339	0%	30	0%
IRC 199 Cited	60	0%	27,752	0%	2,760	0%	-1,175	0%	3,153	0%	3,251	0%	903	0%	136	0%
IRC 263 Cited	29	0%	11,749	0%	126	0%	-56	0%	182	0%	254	0%	87	0%	35	0%
\$10M < Assets < \$100M	29,874	71%	913,975	2%	5,749	1%	-8,461	1%	-2,942	0%	3,138	0%	11,453	5%	437	1%
No Schedule UTP Filed	29,660	70%	899,242	2%	6,381	1%	-8,284	1%	-2,318	0%	3,635	0%	11,094	5%	417	1%
Schedule UTP Filers	214	1%	14,733	0%	-632	0%	-178	0%	-623	0%	-498	0%	359	0%	20	0%
SEC 10K/Public																
All	4,339	10%	37,053,332	73%	830,523	86%	-633,774	95%	856,373	83%	691,661	78%	161,981	71%	65,826	81%
No Schedule UTP Filed	3,109	7%	12,117,934	24%	180,927	19%	-67,760	10%	171,559	17%	128,538	14%	36,666	16%	9,789	12%
Schedule UTP Filers	1,230	3%	24,935,398	49%	649,596	67%	-566,014	85%	684,814	67%	563,123	63%	125,314	55%	56,037	69%
Assets > \$100M	1,176	3%	24,931,339	49%	649,845	67%	-565,835	85%	685,232	67%	563,578	63%	125,261	55%	56,034	69%
IRC 41 Cited	506	1%	7,203,724	14%	259,401	27%	-246,431	37%	257,066	25%	187,595	21%	44,633	20%	12,504	15%
IRC 482 Cited	326	1%	7,798,171	15%	331,638	34%	-459,800	69%	262,031	25%	244,930	28%	46,929	21%	32,918	40%
IRC 162 Cited	164	0%	8,576,916	17%	155,503	16%	-62,374	9%	204,080	20%	149,999	17%	36,415	16%	12,927	16%
IRC 199 Cited	130	0%	2,255,147	4%	192,757	20%	-164,726	25%	148,203	14%	131,989	15%	28,669	13%	14,514	18%
IRC 263 Cited	91	0%	2,304,560	5%	91,407	9%	-63,933	10%	146,988	14%	85,809	10%	16,465	7%	13,329	16%
Assets > \$1B	734	2%	24,734,481	49%	643,929	66%	-562,254	84%	678,196	66%	554,893	62%	122,233	53%	55,550	68%
IRC 41 Cited	241	1%	7,091,740	14%	255,743	26%	-244,305	37%	252,709	25%	182,306	20%	42,867	19%	12,287	15%
IRC 482 Cited	236	1%	7,751,584	15%	329,690	34%	-457,717	68%	260,638	25%	242,481	27%	46,359	20%	32,727	40%
IRC 162 Cited	140	0%	8,565,492	17%	155,244	16%	-62,002	9%	203,803	20%	149,392	17%	36,212	16%	12,905	16%
IRC 199 Cited	93	0%	2,236,003	4%	190,479	20%	-163,783	24%	145,799	14%	129,651	15%	28,007	12%	14,436	18%
IRC 263 Cited	82	0%	2,299,717	5%	91,210	9%	-63,831	10%	146,760	14%	85,625	10%	16,414	7%	13,297	16%

TABLE 2B. UTP Presence by Asset Size and Financial Statement for Tax Year 2012 (Dollar amounts in millions)

	Returns		Total Assets		Worldwide Income (Part I, Line 4)		Nonincludible Foreign Income		Pretax Book		Tax Net Income		Tax After Credits		Foreign Tax Credit	
	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%	Sum	Col%
	SEC 10K/Public Filers of Schedule UTP (continued)															
\$100M < Assets < \$1B	442	1%	196,857	0%	5,917	1%	-3,581	1%	7,035	1%	8,685	1%	3,028	1%	484	1%
IRC 41 Cited	265	1%	111,985	0%	3,658	0%	-2,126	0%	4,357	0%	5,289	1%	1,766	1%	217	0%
IRC 482 Cited	90	0%	46,587	0%	1,949	0%	-2,083	0%	1,393	0%	2,449	0%	570	0%	190	0%
IRC 162 Cited	24	0%	11,424	0%	259	0%	-372	0%	277	0%	606	0%	202	0%	22	0%
IRC 199 Cited	37	0%	19,144	0%	2,278	0%	-943	0%	2,404	0%	2,338	0%	662	0%	77	0%
IRC 263 Cited	9	0%	4,843	0%	198	0%	-102	0%	228	0%	184	0%	51	0%	32	0%
Assets < \$100M	54	0%	4,059	0%	-249	0%	-179	0%	-417	0%	-455	0%	53	0%	3	0%
Audited	17,192	41%	6,933,274	14%	66,555	7%	-24,750	4%	81,973	8%	100,536	11%	36,403	16%	8,336	10%
No Schedule UTP Filed	16,545	39%	4,168,814	8%	52,168	5%	-11,975	2%	61,781	6%	52,981	6%	26,116	11%	2,008	2%
Schedule UTP Filers	647	2%	2,764,460	5%	14,387	2%	-12,775	2%	20,192	2%	47,555	5%	10,287	4%	6,328	8%
Assets > \$100M	533	1%	2,756,837	5%	14,812	2%	-12,765	2%	20,519	2%	47,611	5%	10,097	4%	6,312	8%
Assets > \$1B	174	0%	2,629,659	5%	12,901	1%	-12,173	2%	17,166	2%	42,602	5%	8,229	4%	6,137	8%
\$100M < Assets < \$1B	359	1%	127,178	0%	1,911	0%	-593	0%	3,352	0%	5,009	1%	1,868	1%	175	0%
Assets < \$100M	114	0%	7,624	0%	-425	0%	-10	0%	-327	0%	-56	0%	190	0%	15	0%
Unaudited	20,770	49%	6,910,516	14%	72,433	7%	-10,605	2%	89,533	9%	97,284	11%	30,438	13%	7,187	9%
No Schedule UTP Filed	20,415	48%	4,901,389	10%	44,736	5%	-8,044	1%	55,014	5%	53,769	6%	20,496	9%	4,488	6%
Schedule UTP Filers	355	1%	2,009,126	4%	27,697	3%	-2,561	0%	34,519	3%	43,515	5%	9,943	4%	2,699	3%
Assets > \$100M	309	1%	2,006,076	4%	27,655	3%	-2,573	0%	34,398	3%	43,501	5%	9,828	4%	2,697	3%
Assets > \$1B	171	0%	1,948,429	4%	26,928	3%	-2,012	0%	33,526	3%	42,189	5%	9,224	4%	2,624	3%
\$100M < Assets < \$1B	138	0%	59,647	0%	727	0%	-560	0%	872	0%	1,312	0%	603	0%	74	0%
Assets < \$100M	46	0%	3,050	0%	42	0%	11	0%	121	0%	14	0%	115	0%	1	0%
Non-Public Filers of Schedule UTP	842	2%	4,762,913	9%	42,467	4%	-15,338	2%	54,917	5%	91,112	10%	19,925	9%	9,009	11%
Assets > \$100M	266	1%	887,182	2%	17,097	2%	-1,422	0%	23,521	2%	30,777	3%	8,282	4%	1,664	2%
IRC 482 Cited	228	1%	1,671,483	3%	27,938	3%	-7,324	1%	30,569	3%	32,836	4%	8,079	4%	3,080	4%
IRC 162 Cited	83	0%	1,127,156	2%	9,383	1%	-4,405	1%	9,011	1%	17,670	2%	3,983	2%	2,553	3%
IRC 199 Cited	49	0%	162,052	0%	2,752	0%	-346	0%	4,071	0%	8,292	1%	2,155	1%	195	0%
IRC 263 Cited	37	0%	246,011	0%	4,525	0%	66	0%	6,762	1%	6,543	1%	2,615	1%	19	0%
Assets > \$1B	345	1%	4,576,088	9%	39,829	4%	-14,185	2%	50,692	5%	84,791	10%	17,453	8%	8,761	11%
IRC 41 Cited	95	0%	820,742	2%	16,041	2%	-1,170	0%	21,601	2%	28,477	3%	7,345	3%	1,572	2%
IRC 482 Cited	119	0%	1,626,713	3%	27,530	3%	-6,955	1%	29,670	3%	30,981	3%	7,409	3%	3,027	4%
IRC 162 Cited	50	0%	1,111,835	2%	9,074	1%	-4,395	1%	8,755	1%	17,432	2%	3,847	2%	2,546	3%
IRC 199 Cited	26	0%	153,444	0%	2,270	0%	-114	0%	3,322	0%	7,379	1%	1,914	1%	137	0%
IRC 263 Cited	17	0%	239,105	0%	4,596	0%	21	0%	6,808	1%	6,473	1%	2,580	1%	16	0%
\$100M < Assets < \$1B	497	1%	186,825	0%	2,638	0%	-1,153	0%	4,225	0%	6,321	1%	2,471	1%	248	0%
IRC 41 Cited	171	0%	66,441	0%	1,056	0%	-252	0%	1,920	0%	2,300	0%	937	0%	92	0%
IRC 482 Cited	109	0%	44,769	0%	408	0%	-370	0%	899	0%	1,855	0%	670	0%	53	0%
IRC 162 Cited	33	0%	15,321	0%	309	0%	-10	0%	256	0%	238	0%	137	0%	7	0%
IRC 199 Cited	23	0%	8,608	0%	482	0%	-232	0%	749	0%	913	0%	241	0%	58	0%
IRC 263 Cited	20	0%	6,905	0%	-71	0%	46	0%	-46	0%	70	0%	36	0%	3	0%

Figure 1 summarizes all Form 1120 corporate filers and nonfilers of Schedule UTP with assets of \$10 million or more for both 2011 and 2012. This figure shows that a minority percentage of Form 1120 corporate taxpayers filed a Schedule UTP for both years. However, this minority percentage of filers represents a majority of the percentage of total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit.

Figure 2A compares the percentages of Form 1120 corporate taxpayers with assets of \$10 million or more that filed a Schedule UTP in 2011 and 2012 to the Form 1120 corporate taxpayers with assets of \$1 billion or more that filed a Schedule UTP for those same years. Figure 2A shows that Form 1120 corporate taxpayers with assets of \$1 billion or more comprise most of the percentages of total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for all taxpayers with assets of \$10 million or more that filed a Schedule UTP in 2011 and 2012.

Figure 2B compares the percentages of Form 1120 corporate taxpayers with assets of \$10 million or more that filed a Schedule UTP in 2011 and 2012 to the Form 1120 corporate taxpayers with SEC financial statements and assets of \$10 million or more that filed a Schedule UTP for those same years. Figure 2B shows that Form 1120 corporate taxpayers with SEC financial statements and assets of \$10 million or more make up most of the percentages of total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for all taxpayers with assets of \$10 million or more that filed a Schedule UTP for 2011 and 2012.

Figure 2C contrasts the percentages of Form 1120 corporate taxpayers with assets of \$10 million or more that filed a Schedule UTP in 2011 and 2012 to the Form 1120 corporate taxpayers with SEC financial statements and assets of \$1 billion or more that filed a Schedule UTP for those same years. This figure shows that Form 1120 corporate taxpayers with SEC financial statements and assets of \$1 billion or more make up most of the percentages of total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for all taxpayers with assets of \$10 million or more that filed a Schedule UTP for 2011 and 2012.

Details for the 2011 LB&I corporation filers and nonfilers of Schedule UTP show:

- Only about 5 percent of the total LB&I Form 1120 population filed a Schedule UTP (2,160 taxpayers). However, this small minority represents 70 percent of the total assets of the LB&I Form 1120 population, 94 percent of the worldwide income, 90 percent of the nonincludible foreign income, 91 percent of the pretax book income, 85 percent of the tax net income, 66 percent of the tax after credits, and 85 percent of the foreign tax credit.
- Of the 2,160 taxpayers that filed a Schedule UTP (5 percent of the total returns), 2,074 of those returns were for taxpayers with \$100 million or more in assets. These 2,074 taxpayers represent the exact same percentages reported in the bullet point above. Therefore, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those LB&I Form 1120 taxpayers with assets under \$100 million are negligible in aggregate compared to the aggregate amounts for those LB&I Form 1120 taxpayers with assets of \$100 million or more.

Highlights of the 2011 Form 1120 corporation Schedule UTP filers by asset size show:

- 1,072 taxpayers (3 percent of the total returns) with assets of \$1 billion or greater filed a Schedule UTP. This group of LB&I Form 1120 taxpayers is significant as they represent 69 percent of the total assets, 92 percent of the worldwide income, 89 percent of the nonincludible foreign income, 90 percent of the pretax book income, 84 percent of the tax net income, 63 percent of the tax after credits, and 84 percent of the foreign tax credit. The percentages for these 1,072 taxpayers with assets of \$1 billion or greater are almost identical to those percentages in the first bullet point of the section above for the 2,160 taxpayers that filed a Schedule UTP. In other words, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those LB&I Form 1120 taxpayers with assets of \$100 million to \$1 billion are small in aggregate compared to the aggregate amounts for those LB&I Form 1120 taxpayers with assets of \$1 billion or more.

FIGURE 1. 2011–2012 U.S. Corporation Schedule UTP Filers and Nonfilers (Assets of \$10M or More): Percentages of Returns/Assets/Income Categories/Tax After Credits

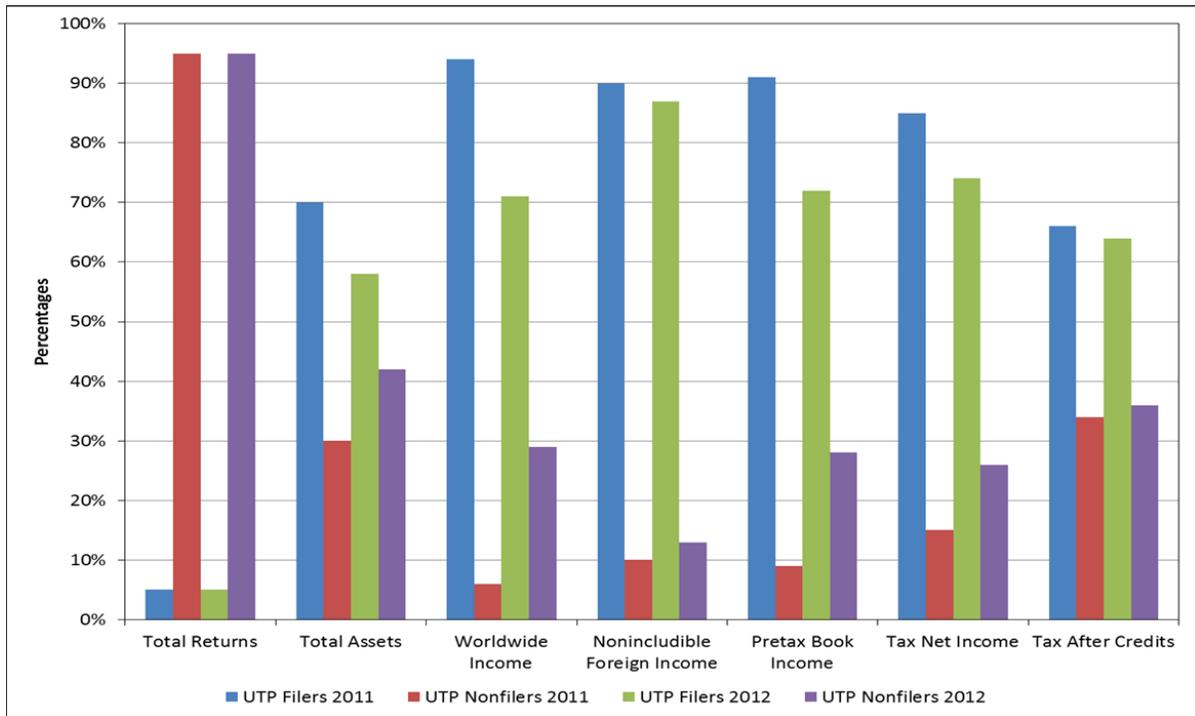


FIGURE 2A. 2011–2012 U.S. Corporation Schedule UTP Filers with Assets of \$10M or More and Filers with Assets of \$1B or More: Percentages of Returns/Assets/Income Categories/Tax After Credits

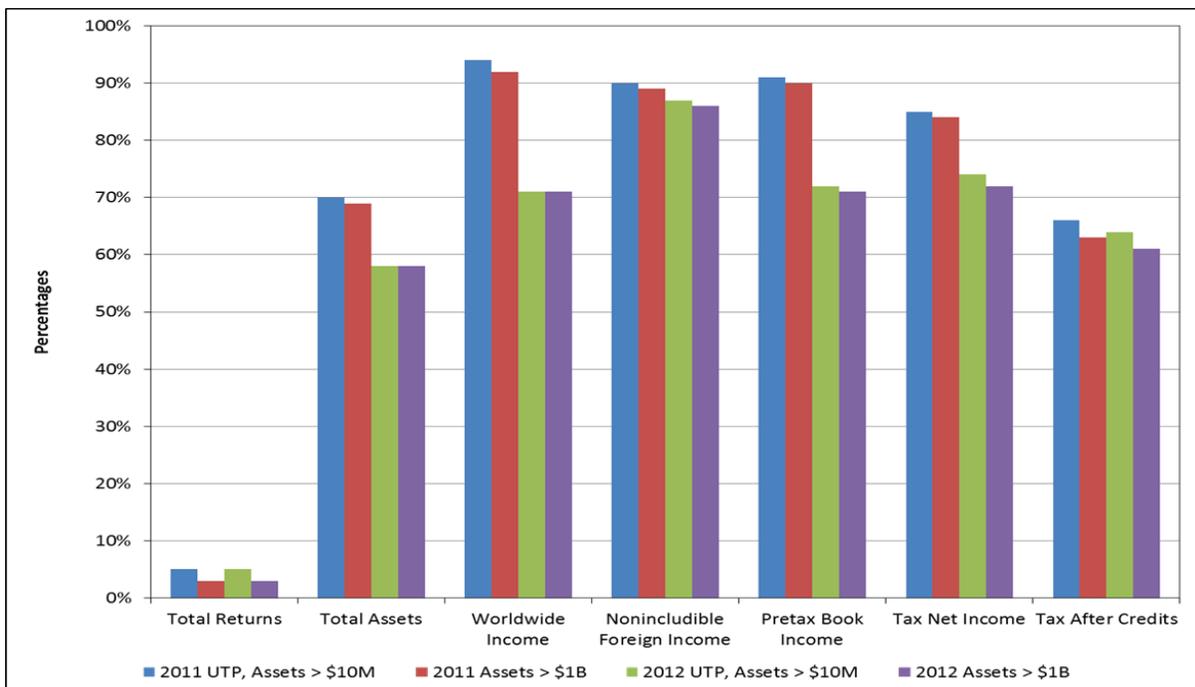


FIGURE 2B. 2011–2012 U.S. Corporation Schedule UTP Filers with Assets of \$10M or More and SEC Filers with Assets of \$10M or More: Percentages of Returns/Assets/Income Categories/Tax After Credits

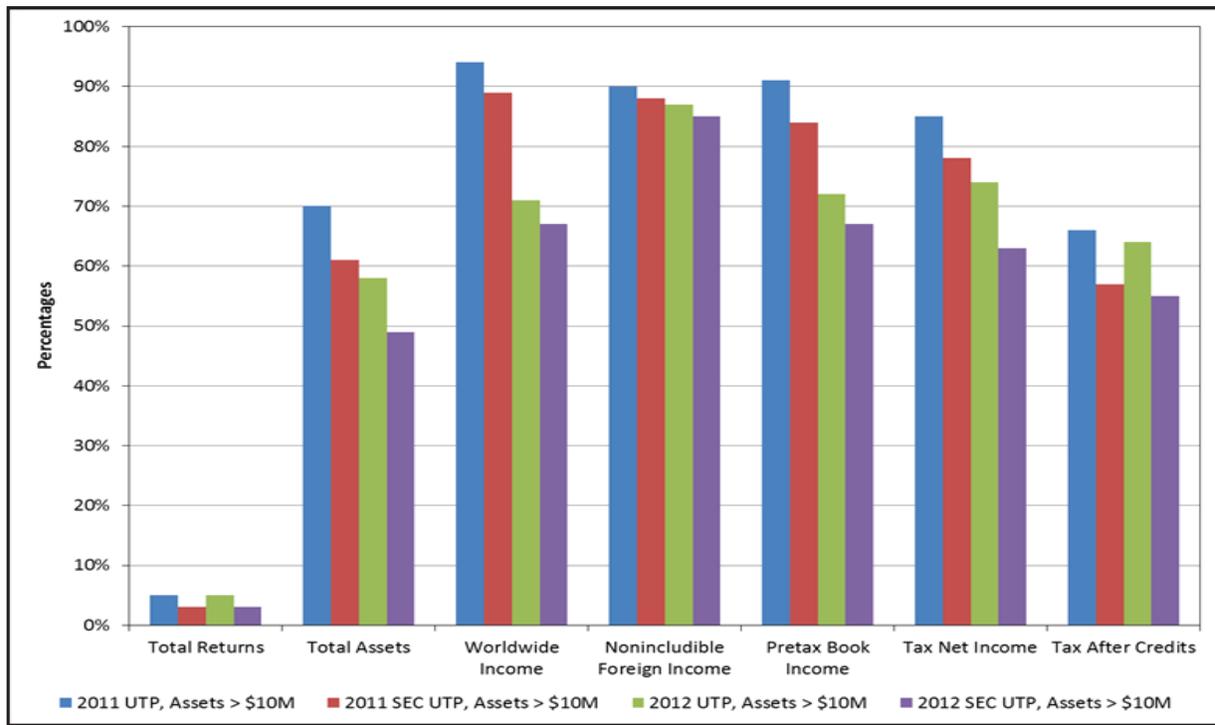
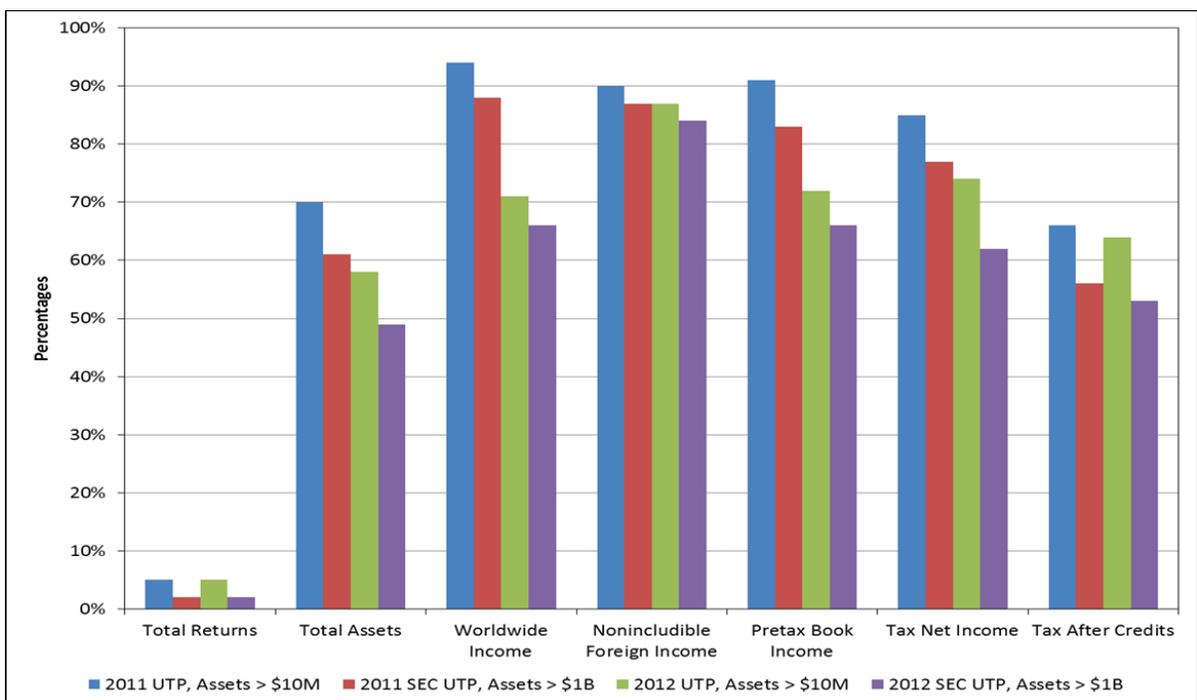


FIGURE 2C. 2011–2012 U.S. Corporation Schedule UTP Filers with Assets of \$10M or More and SEC Filers with Assets of \$1B or More: Percentages of Returns/Assets/Income Categories/Tax After Credits



- For LB&I Form 1120 taxpayers with assets between \$100 million and \$1 billion, 1,002 (2 percent of the total returns) filed a Schedule UTP. This represents 1 percent of the total assets, 2 percent of the worldwide income, 1 percent of the nonincludible foreign income, 1 percent of the pretax book income, 2 percent of the tax net income, 3 percent of the tax after credits, and 1 percent of the foreign tax credit.

There are several key observations about the 2011 Form 1120 corporation Schedule UTP filers by financial statement type:

- Of the 4,488 taxpayers that reported filing a 10K with the SEC (SEC 10K/Public), 1,238 (3 percent of the total returns) filed a Schedule UTP. This signifies another important group of taxpayers, as they represent 61 percent of the total assets of LB&I Form 1120 taxpayers, 89 percent of the worldwide income, 88 percent of the nonincludible foreign income, 84 percent of the pretax book income, 78 percent of the tax net income, 57 percent of the tax after credits, and 79 percent of the foreign tax credit.
- Of the 1,238 SEC 10K/Public taxpayers that filed a Schedule UTP as noted in the bullet above, 1,227 (3 percent of the total returns) were SEC 10K/Public taxpayers with \$100 million or more in assets. These 1,227 taxpayers represent the exact same percentages reported in the bullet point above. Therefore, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those LB&I Form 1120 taxpayers with assets under \$100 million are negligible in aggregate compared to the aggregate amounts for those LB&I taxpayers with assets of \$100 million or more.
- Looking further into the 1,238 SEC 10K/Public taxpayers that filed a Schedule UTP (see the first bullet point of this section), 748 of these taxpayers (2 percent of the total returns) have assets of \$1 billion or greater. This group of LB&I Form 1120 taxpayers is significant, as they represent almost the exact same percentages noted in the first bullet point of this section for all SEC 10K/Public filers since these 748 taxpayers report 61 percent of the total assets, 88 percent of the worldwide income, 87 percent of the nonincludible foreign income, 83 percent of the pretax book income, 77 percent of the tax net income, 56 percent of the tax after credits, and 79 percent of the foreign tax credit. In other words, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those SEC 10K/public taxpayers with assets of \$100 million to \$1 billion are small in aggregate compared to the aggregate amounts for those SEC 10K/Public taxpayers with assets of \$1 billion or more.

Finally, a review of the Non-Public taxpayers (with Audited or Unaudited financial statements) that filed a Schedule UTP in 2011 shows that 846 Non-Public taxpayers with assets greater than \$100 million filed a Schedule UTP. This represents 2 percent of the total LB&I Form 1120 returns, 9 percent of the total assets, 5 percent of the worldwide income, 2 percent of the nonincludible foreign income, 7 percent of the pretax book income, 7 percent of the tax net income, 9 percent of the tax after credits, and 5 percent of the foreign tax credit.

Analysis of the 2012 LB&I Form 1120 corporation filers and nonfilers of Schedule UTP shows:

- Again, only about 5 percent of the total LB&I population filed a Schedule UTP (2,232 taxpayers). However, this small minority represents 58 percent of the total assets of LB&I Form 1120 taxpayers (compared to 70 percent in 2011), 71 percent of the worldwide income (compared to 94 percent in 2011), 87 percent of the nonincludible foreign income (compared to 90 percent in 2011), 72 percent of the pretax book income (compared to 91 percent in 2011), 74 percent of the tax net income (compared to 85 percent in 2011), 64 percent of the tax after credits (compared to 66 percent in 2011), and 80 percent of the foreign tax credit (compared to 85 percent in 2011).
- As discussed in the next two bullets, the number of SEC 10K/Public Schedule UTP filers with \$100 million or more in assets decreased a net 51 returns from 2011 to 2012 and the total assets of such filers decreased a net \$6.4 trillion. A small number of very large corporations with SEC 10K/Public financial statements that filed Schedule UTP in 2011, but did not file Schedule UTP in 2012, decreased the aggregate percentage dollar amounts reported in 2012 by Schedule UTP filers as shown in the prior bullet.
- In 2011, Schedule UTP filers included 1,227 with SEC 10K/Public financial statements and \$100 million or more in assets reporting \$31.3 trillion in assets, 61 percent of total assets of all corporations with \$10 million or more assets filing Form 1120.

- In 2012, Schedule UTP filers included 1,176 with SEC 10K/Public financial statements and \$100 million or more in assets reporting \$24.9 trillion in assets, 49 percent of total assets of all corporations with \$10 million or more in assets filing Form 1120.
- Of the 2,232 LB&I Form 1120 taxpayers that filed a Schedule UTP in 2012 (5 percent of the total returns), 2,018 were taxpayers with \$100 million or more in assets. These 2,018 taxpayers again represent the exact same percentages reported in the bullet point above except for tax after credits (63 percent were filers with assets of \$100 million or greater rather than the 64 percent for all Schedule UTP filers). Again in 2012, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those LB&I Form 1120 taxpayers with assets under \$100 million are negligible in aggregate compared to the aggregate amounts for those LB&I Form 1120 taxpayers with assets of \$100 million or more.

There are several key observations about the 2012 LB&I Form 1120 corporation Schedule UTP filers by asset size:

- 1,079 taxpayers in 2012 (3 percent of the total returns and almost the exact same number of 1,072 taxpayers for 2011) with assets of \$1 billion or greater filed a Schedule UTP. This group of taxpayers is significant, as they represent 58 percent of the total assets of LB&I Form 1120 taxpayers (compared to 69 percent for 2011), 71 percent of the worldwide income (compared to 92 percent for 2011), 86 percent of the nonincludible foreign income (compared to 89 percent for 2011), 71 percent of the pretax book income (compared to 90 percent for 2011), 72 percent of the tax net income (compared to 84 percent for 2011), 61 percent of the tax after credits (compared to 63 percent for 2011), and 79 percent of the foreign tax credit (compared to 84 percent for 2011). The percentages for these 1,079 taxpayers with assets of \$1 billion or greater are almost identical to those percentages in the first bullet point of the section above for the 2,232 taxpayers that filed a Schedule UTP. In other words, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those LB&I Form 1120 taxpayers with assets of \$100 million to \$1 billion are small in aggregate compared to the aggregate amounts for those LB&I Form 1120 taxpayers with assets of \$1 billion or more.
- For taxpayers with assets between \$100 million and \$1 billion, 939 (2 percent of the total returns) filed a Schedule UTP in 2012. This represents 1 percent of the total assets, 1 percent of the worldwide income, 1 percent of the nonincludible foreign income, 1 percent of the pretax book income, 2 percent of the tax net income, 2 percent of the tax after credits, and 1 percent of the foreign tax credit (almost identical percentages to 2011).

Analysis of the 2012 LB&I Form 1120 Schedule UTP filers by financial statement type shows:

- Of the 4,339 taxpayers in 2012 (4,488 in 2011) that reported filing a 10K with the SEC (SEC 10K/Public), 1,230 (3 percent of the total returns) filed a Schedule UTP (compared to 1,238 in 2011). This signifies another important group of taxpayers, as they represent 49 percent of the LB&I Form 1120 taxpayer total assets (compared to 61 percent in 2011), 67 percent of the worldwide income (compared to 89 percent in 2011), 85 percent of the nonincludible foreign income (compared to 88 percent in 2011), 67 percent of the pretax book income (compared to 84 percent in 2011), 63 percent of the tax net income (compared to 78 percent in 2011), 55 percent of the tax after credits (compared to 57 percent in 2011), and 69 percent of the foreign tax credit (compared to 79 percent in 2011).
- Of the 1,230 SEC 10K/Public taxpayers that filed a Schedule UTP for 2012 as noted in the bullet above, 1,176 of those returns (3 percent of the total returns) were for SEC 10K/Public taxpayers with \$100 million or more in assets. These 1,176 taxpayers represent the exact same percentages reported in the bullet point above. Again this year, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those LB&I Form 1120 taxpayers with assets under \$100 million are negligible in aggregate compared to the aggregate amounts for those LB&I taxpayers with assets of \$100 million or more.
- Looking further into the 1,230 SEC 10K/Public taxpayers that filed a Schedule UTP (see the first bullet point of this section for 2012), 734 of these taxpayers (2 percent of the total returns) have assets of

\$1 billion or greater (compared to 748 taxpayers for 2011). As with 2011, this group of taxpayers is significant, as they represent almost the exact same percentages noted in the first bullet point of this section for all SEC 10K/Public filers since these 734 taxpayers report 49 percent of the total assets, 66 percent of the worldwide income, 84 percent of the nonincludible foreign income, 66 percent of the pretax book income, 62 percent of the tax net income, 53 percent of the tax after credits, and 68 percent of the foreign tax credit. The percentages for these 734 SEC 10K/Public taxpayers that file a Schedule UTP are once again also very close to the percentages noted in the first bullet point under the section for 2012 detailing the filers and nonfilers. In other words, total assets, worldwide income, nonincludible foreign income, pretax book income, tax net income, tax after credits, and foreign tax credit for those SEC 10K/public taxpayers with assets of \$100 million to \$1 billion are small in aggregate compared to the aggregate amounts for those SEC 10K/Public taxpayers with assets of \$1 billion or more.

Lastly, the review of the Non-Public taxpayers that filed a Schedule UTP in 2012 shows that 842 Non-Public taxpayers with assets greater than \$100 million filed a Schedule UTP in 2012 (compared to 846 in 2011). This represents 2 percent of the total returns, 9 percent of the total assets, 4 percent of the worldwide income, 2 percent of the nonincludible foreign income, 5 percent of the pretax book income, 10 percent of the tax net income, 9 percent of the tax after credits, and 11 percent of the foreign tax credit (almost the exact same percentages as 2011 in all categories except foreign tax credit was 5 percent in 2011).

Overall Comments:

- The top two IRC sections cited in both 2011 and 2012 are Section 41 (Credit for increasing research activities) and Section 482 (Allocation of income and deductions among taxpayers—“Transfer Pricing”). These IRC sections represent large calculations that occur annually for large multi-national taxpayers. It is unlikely that the entire amounts reported on the filed tax returns are uncertain. Rather, the uncertainty probably exists due to the complexity of these calculations and the methodologies used to compute the amounts reported on the returns. If examined, the IRS may contend the methodologies used to calculate the amounts should be changed.
- The SEC 10K/Public sub-population with assets of \$100M or more essentially account for all of the dollar amounts for the Form 1120 population with assets of \$10 million or more. For both 2011 and 2012, a minority of taxpayers with \$100 million or more in assets and an SEC 10K/Public financial statement filed a Schedule UTP. This is the group that will be examined in detail in Part III of this study.
- The number of SEC 10K/Public Schedule UTP filers with total assets of \$100 million or more decreased a net 51 returns from 2011 to 2012 and the total assets of such filers decrease a net \$6.4 trillion. A small number of very large corporations with SEC 10K/Public financial statements that filed Schedule UTP in 2011 but did not file Schedule UTP in 2012 decrease the aggregate percentage dollar amounts reported in 2012 by Schedule UTP filers.

Part III. 2011–2012 Analysis of M-3 Profiles

1. UTP Filers vs Nonfilers

The balance of this study will focus on data describing characteristics of Schedule UTP filers and nonfilers with SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more but with Audited (but not SEC 10K/Public) or Unaudited financial statements.³⁹

Tables 3A and 3B and Figures 3A and 3B present 2012 data describing characteristics of Schedule UTP filers and nonfilers with SEC 10K/Public financial statements and \$100 million or more in assets. Figures 3A and 3B include 2011 data for comparison. Figure 3C presents 2011 and 2012 average data.

³⁹ A Schedule UTP filer may file because it is required to file or because it files voluntarily. A Schedule UTP nonfiler may be a nonfiler either because it is not required to file or because it is required to file but failed to file. For tax years beginning in 2012 and later, the asset threshold for Schedule UTP drops to \$50 million or more. This study does not include the new asset range of \$50 million to \$100 million for Schedule UTP in the detailed analysis in Part III because a 2-year comparison is not possible.

TABLE 3A. 2012 U.S. Corporation 1120 Schedule M-3: Mini M-3 by UTP Status: SEC 10K/Public: Assets \$100 Million or More

	\$ Million				Percentage of Adjusted Total Income Book				Percentage of Pretax Book				
	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Total Difference	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Total Difference	Column B Temporary	Column C Permanent	Total Difference
UTP Filers													
Adj. Other income—no difference	6,615,611	0	0	6,615,611	0	179.17%	0.00%	0.00%	179.17%	0.00%	0.00%	0.00%	0.00%
Adjusted COGS	-4,071,080	-7,006	-201	-4,078,286	-7,208	-110.25%	-0.19%	-0.01%	-110.45%	-0.20%	-1.02%	-0.03%	-1.05%
Adjusted gross profit	2,544,531	-7,006	-201	2,537,325	-7,208	68.91%	-0.19%	-0.01%	68.72%	-0.20%	-1.02%	-0.03%	-1.05%
Specified income	999,311	29,906	-32,262	996,825	-2,356	27.06%	0.81%	-0.87%	27.00%	-0.06%	4.36%	-4.71%	-0.34%
Adj. Other income—difference	148,617	24,110	-16,732	155,995	7,378	4.02%	0.65%	-0.45%	4.22%	0.20%	3.52%	-2.44%	1.08%
Adjusted total income	3,692,459	47,010	-49,195	3,690,145	-2,186	100.00%	1.27%	-1.33%	99.94%	-0.06%	6.86%	-7.18%	-0.32%
Specified expense/deduction	-987,005	-35,760	8,245	-1,014,390	-27,515	-26.73%	-0.97%	0.22%	-27.47%	-0.75%	-5.22%	1.20%	-4.02%
Other exp./ded.—difference	-1,090,550	-40,776	119	-1,131,208	-40,657	-29.53%	-1.10%	0.00%	-30.64%	-1.10%	-5.95%	0.02%	-5.93%
Other exp./ded.—no difference	-929,673	0	0	-929,673	0	-25.18%	0.00%	0.00%	-25.18%	0.00%	0.00%	0.00%	0.00%
Pretax net income	685,231	-29,526	-40,831	614,874	-70,358	18.56%	-0.80%	-1.11%	16.65%	-1.91%	-4.31%	-5.96%	-10.27%
UTP Nonfilers													
Adj. Other income—no difference	2,365,953	0	0	2,365,959	0	157.66%	0.00%	0.00%	157.66%	0.00%	0.00%	0.00%	0.00%
Adjusted COGS	-1,372,691	-15,165	1,927	-1,385,952	-13,238	-91.47%	-1.01%	0.13%	-92.36%	-0.88%	-8.57%	1.09%	-7.48%
Adjusted gross profit	993,262	-15,165	1,927	980,007	-13,238	66.19%	-1.01%	0.13%	65.31%	-0.88%	-8.57%	1.09%	-7.48%
Specified income	440,252	-89,400	13,397	364,261	-76,003	29.34%	-5.96%	0.89%	24.27%	-5.06%	-50.54%	7.57%	-42.96%
Adj. Other income—difference	67,118	11,920	-3,818	75,222	8,103	4.47%	0.79%	-0.25%	5.01%	0.54%	6.74%	-2.16%	4.58%
Adjusted total income	1,500,632	-92,645	11,506	1,419,490	-81,138	100.00%	-6.17%	0.77%	94.59%	-5.41%	-52.37%	6.50%	-45.86%
Specified expense/deduction	-493,009	61,604	-708	-432,107	60,896	-32.85%	4.11%	-0.05%	-28.80%	4.06%	34.82%	-0.40%	34.42%
Other exp./ded.—difference	-402,488	-19,387	1,339	-420,537	-18,048	-26.82%	-1.29%	0.09%	-28.02%	-1.20%	-10.96%	0.76%	-10.20%
Other exp./ded.—no difference	-428,228	0	0	-428,228	0	-28.54%	0.00%	0.00%	-28.54%	0.00%	0.00%	0.00%	0.00%
Pretax net income	176,907	-50,428	12,137	138,618	-38,290	11.79%	-3.36%	0.81%	9.24%	-2.55%	-28.51%	6.86%	-21.64%

TABLE 3B. 2012 U.S. Corporation 1120 Schedule M-3: M-3, Part I, by UTP Status: SEC 10K/ Public: Assets \$100 Million or More

	Mean Amount Reported (\$ Million)		
	UTP Filer	UTP Nonfiler	Difference
Worldwide financial net income	553	87	466
(Foreign nonincludible income)	-481	-31	-450
(U.S. nonincludible income)	-32	-8	-23
Other includible income	1	-1	2
Adjustments to eliminations	294	5	289
Other adjustments	131	12	119
Book net income includible corp.	472	64	408
Pretax net income—book	583	84	499
Tax net income	523	66	457
Total BTD Difference	-60	-18	-42
Total assets—Schedule L	21,200	5,719	15,481
Assets—Financial statements	18,980	3,018	15,962
Assets—Foreign nonincludible	7,571	530	7,041
Assets—U.S. nonincludible	653	233	419
Assets—Other includible	118	18	100
	Percentage of Adjusted Total Income Book		
	UTP Filer	UTP Nonfiler	Difference
Worldwide financial net income	17.60%	12.21%	5.39%
(Foreign nonincludible income)	-15.32%	-4.40%	-10.92%
(U.S. nonincludible income)	-1.01%	-1.15%	0.15%
Other includible income	0.03%	-0.11%	0.14%
Adjustments to eliminations	9.37%	0.72%	8.66%
Other adjustments	4.16%	1.69%	2.47%
Book net income includible corp.	15.02%	9.00%	6.03%
Pretax net income—book	18.56%	11.79%	6.77%
Tax net income	16.65%	9.24%	7.41%
Total BTD difference	-1.91%	-2.55%	0.65%
	Distribution of Returns		
	UTP Filer	UTP Nonfiler	Total
Number of Returns	1,176	2,112	3,288
Share of Returns	35.8%	64.2%	100.0%

Schedule UTP filers are generally larger than nonfilers when study is restricted to the 3,288 Form 1120 corporate returns in 2012 with SEC 10K/Public financial statements and \$100 million or more in assets. The mean asset size as reported on the Form 1120 Schedule L by such Schedule UTP filers and nonfilers is \$21,200 million for the 1,176 filers and \$5,719 million for the 2,112 nonfilers (see Table 3B). The mean worldwide financial net income of the filers (reported on Schedule M-3, Part I, line 4) is \$553 million compared to \$87 million for the nonfilers. The mean foreign nonincludible income of filers (reported on Schedule M-3, Part I, line 5) is -\$481 million (shown as negative since the income is removed in calculating book income) compared to -\$31 million for nonfilers. After the required Schedule M-3, Part I, adjustments, the mean book income for filers (reported on Schedule M-3, Part I, line 11) is \$472 million compared to \$64 million for the nonfilers. Adding back U.S. Federal tax expense, mean pretax book income is \$583 million for the filers compared to \$84 million for the nonfilers. Filers introduce mean BTD of -\$60 million to adjust pretax book to mean tax net income (reported on Schedule M-3, Part II, line 30, column D, and on Form 1120 page 1, line 28) of \$523 million compared to nonfilers that introduce mean BTD of -\$18 million to adjust pretax book to mean tax net income of \$66 million.⁴⁰

⁴⁰ Negative total BTD adjustments added to pretax book income result in tax net income that is lower than pretax book income.

FIGURE 3A. 2011–2012 U.S. Corporation Key M-3 Data as Percentage of Adjusted Total Book Income for SEC/10K Public Financial Statements by UTP Filing Status

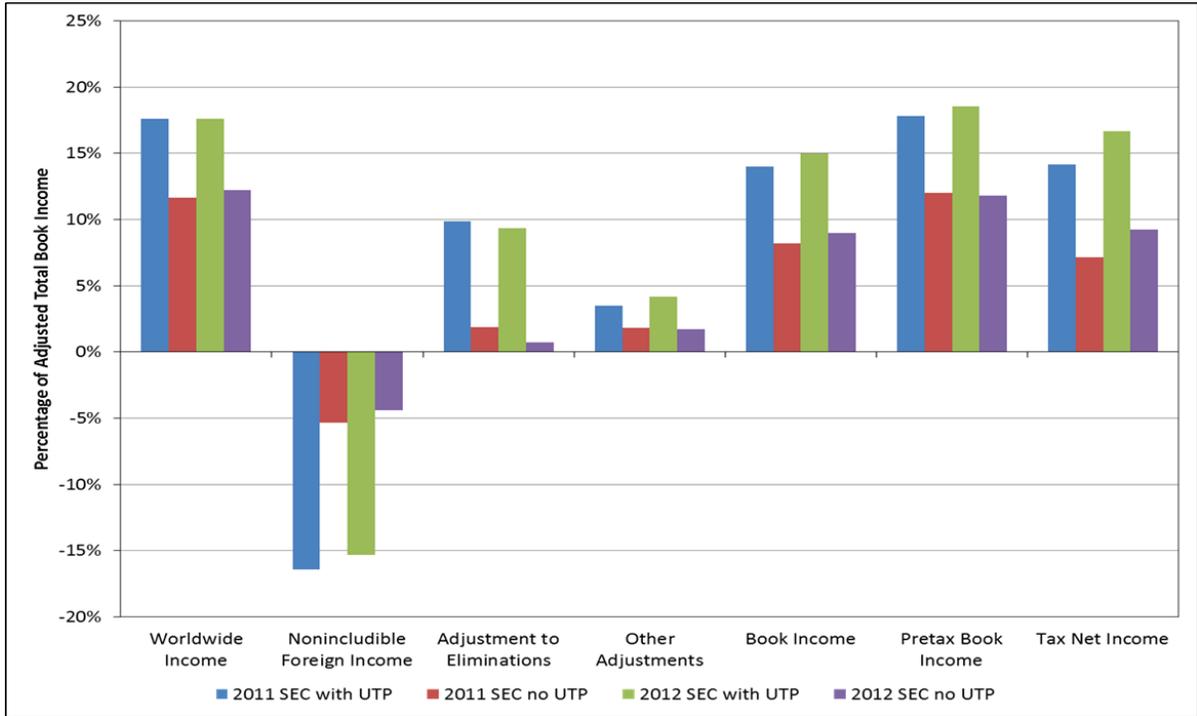


FIGURE 3B. 2011–2012 U.S. Corporation M-3: BTD as Percentage of Pretax Book Income for SEC 10K/Public Financial Statements by UTP Filing Status

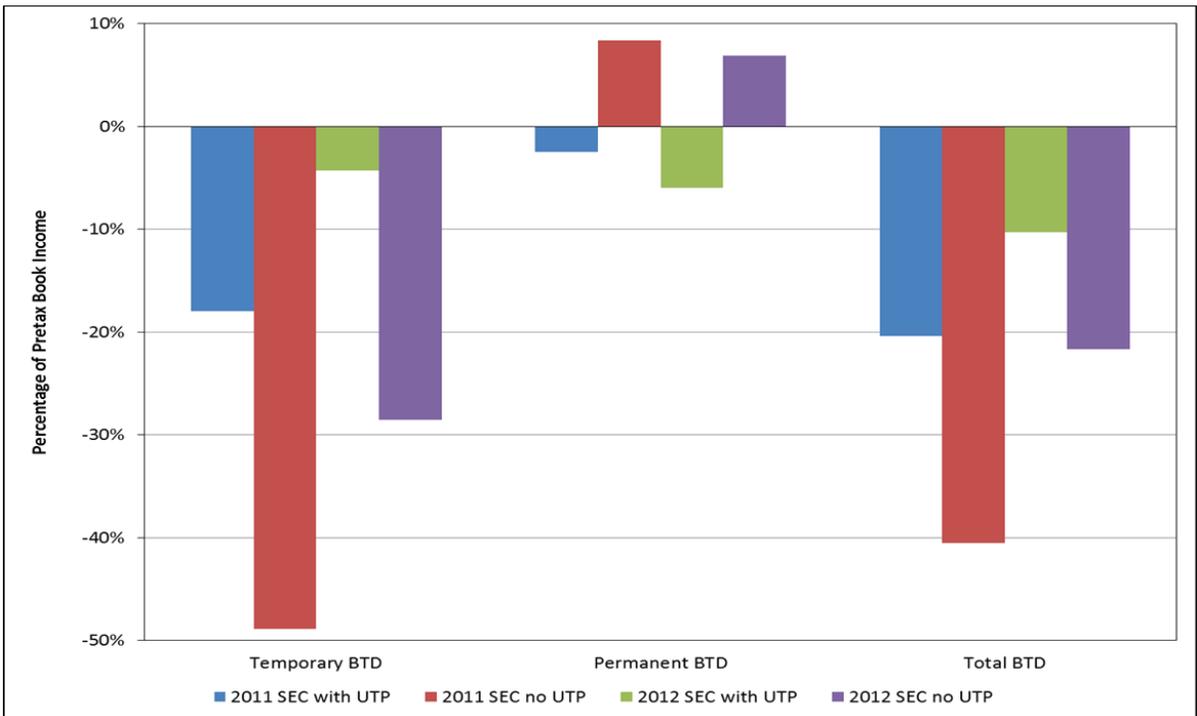
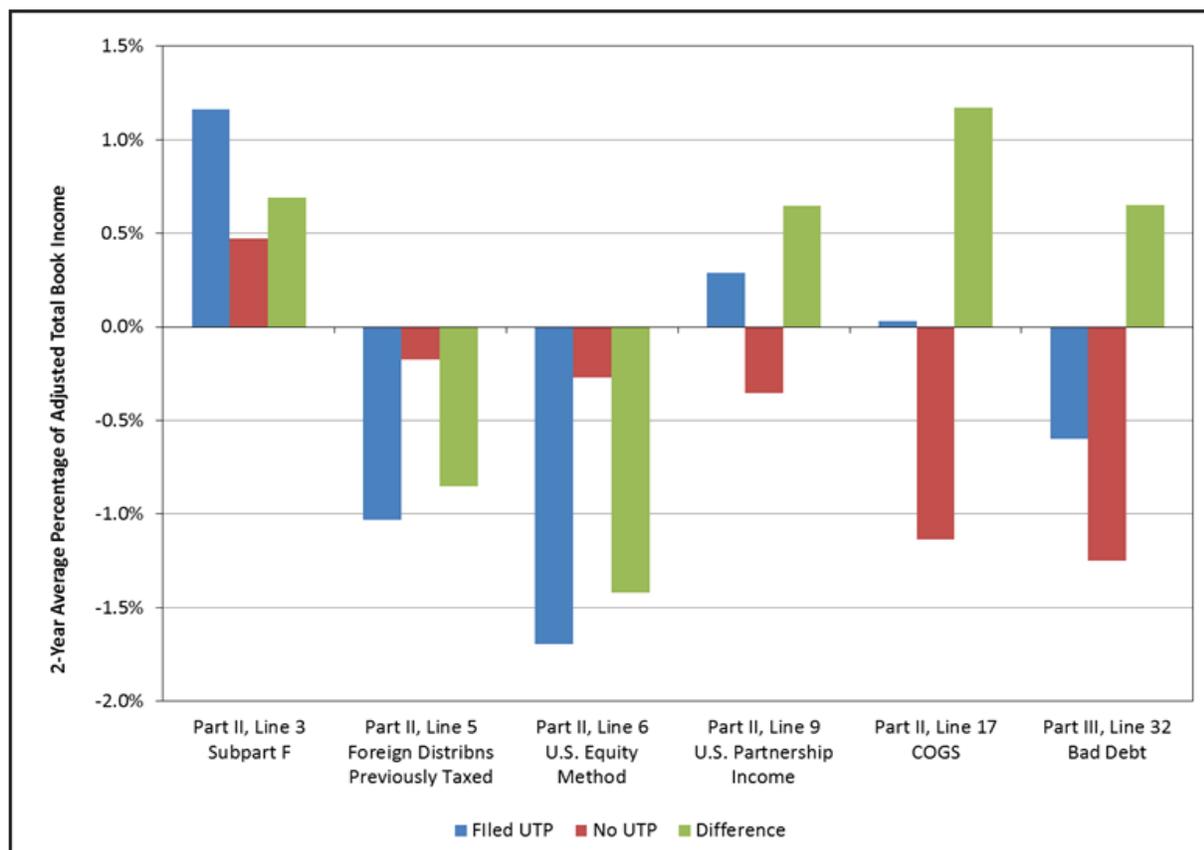


FIGURE 3C. 2011–2012 U.S. Corporation M-3: 2-Year Average Total BTD as Percentage of Adjusted Total Book Income for Top 6 Lines—UTP Filers and Nonfilers (SEC/10K Public Financial Statements)



We scaled our initial aggregate dollar data by the adjusted-total-income measure we developed for our Mini M-3 analysis to make data for filers and nonfilers more comparable. We used the adjusted-total-income book amount as a common-size scaling factor and compared percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers and nonfilers, total pretax income BTD is expressed as a percentage of total pretax book income (see Tables 3A and 3B).

Table 3B shows that Schedule UTP filers in 2012 having SEC 10K/Public financial statements and \$100 million or more in assets report:

- more worldwide income scaled as a percentage of adjusted total income than similar nonfilers (17.60 percent compared to 12.21 percent);
- more scaled foreign nonincludible income (-15.32 percent versus -4.40 percent);
- more scaled book income (15.02 percent versus 9.00 percent);
- more scaled pretax book income (18.56 percent versus 11.79 percent);
- more scaled tax net income (16.65 percent versus 9.24 percent); but
- *less* negative scaled BTB (-1.91 percent versus -2.55 percent).

If the BTD is scaled by pretax book income, the filers reduce pretax book income by -10.27 percent to determine tax net income compared to a reduction of -21.64 percent by nonfilers (see Table 3A, last row, last column, in each panel).⁴¹

Figure 3A compares Schedule M-3 data as a percentage of adjusted total income for 2011 and 2012 Schedule UTP filers and nonfilers with SEC 10K/Public financial statements and \$100 million or more in assets for seven key items: worldwide financial income (Part I, line 4); nonincludible foreign income (Part I, line 5); adjustments to eliminations (Part I, line 8); other adjustments (Part I, line 10); book income (Part I, line 11); pretax book income; and tax net income. For both years filers report larger scaled amounts for the seven items than nonfilers.

Figure 3B uses pretax book income for scaling 2011 and 2012 BTD, temporary, permanent, and total, for Schedule UTP filers and nonfilers with SEC 10K/Public financial statements and \$100 million or more in assets. Filers have less negative temporary BTD than nonfilers, but more negative permanent BTD. In aggregate, the nonfilers report positive permanent BTD partially offsetting their large negative temporary BTD. Filers have less negative total BTD than nonfilers even after the offsetting effects of the permanent BTD for the nonfilers. Stated differently, the otherwise similar nonfilers use negative total BTD to reduce tax net income *more* than the filers.

In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

Figure 3C reports total BTD for six specific Schedule M-3 lines for Schedule UTP filers and nonfilers with SEC 10K/Public financial statements and \$100 million or more in assets. The six lines are selected from and represent approximately 10 percent of the 68 specified and other-with-difference lines on Schedule M-3, Parts II and III.⁴² The total BTD for each of the six lines is expressed as a percentage of adjusted total income and averaged for 2011 and 2012. The lines are the top (that is, the most extreme) six lines on Schedule M-3, Parts II and III, in terms of the absolute value of the difference in 2-year-average scaled BTD between filers and nonfilers.⁴³ The percentages for each of these lines are provided below.

Schedule M-3 Lines Featured in Figure 3C		Filers	Nonfilers	Difference
Part II, line 3:	Inclusion in tax income of subpart F foreign income	1.17%	0.47%	0.69%
Part II, line 5:	Exclusion in tax income of previously taxed foreign distributions	-1.03%	-0.18%	-0.85%
Part II, line 6:	Exclusion from tax income of U.S. equity method income	-1.69%	-0.27%	-1.42%
Part II, line 9:	Adjustments to U.S. partnership income to include all Schedule K-1 income in tax income	0.29%	-0.36%	0.65%
Part II, line 17:	Adjustments to COGS in tax income	0.03%	-1.14%	1.17%
Part III, line 32:	Adjustments to bad debt expense/deduction recognition in tax income	-0.60%	-1.25%	0.65%

⁴¹ A similar pattern of BTD reductions for Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets occurred in 2010 and 2011. See Boynton, DeFilippes, Legel, and Rupert (2014). See Towery (2015): “My results suggest firms found ways to change their financial reporting for tax uncertainty to avoid disclosing positions unknown to the IRS on Schedule UTP. Specifically, I find that although firms decrease financial statement reserves for tax uncertainty in response to Schedule UTP, firms continue claiming uncertain tax positions with the adoption of Schedule UTP. Overall, my results imply that linking tax return disclosures of uncertain tax positions to financial reporting for tax uncertainty can distort financial reporting disclosures of tax uncertainty.”

⁴² There are 31 lines on Part II and 37 on Part III for a total of 68, but that is effectively reduced to 60 for our study. We exclude Federal tax expense (Part II, lines 1 and 2) because it is not part of pretax book. We combine, that is, net, BTD reported for asset dispositions on the multiple lines of Part II, lines 23a through 23g, because corporations use line 23a to reverse out all book income for asset disposition and use lines 23b through 23g to bring in the tax income effects.

⁴³ For the current study, we exclude interest income (Part II, line 13) and interest expense (Part III, line 8) as key lines because of a data anomaly in the 2011 and 2012 data. A number of corporations with SEC 10K/Public financial statements and \$100 million or more in assets in both 2011 and 2012 were Schedule UTP 2011 filers but 2012 nonfilers. This group of 2011 filers and 2012 nonfilers includes corporations reporting large temporary BTD in *both* 2011 and 2012 on *both* interest income and interest expense. The temporary BTD in 2012 for interest income and interest expense are particularly large when the corporations are nonfilers. Assuming the taxpayers made the correct decisions to file Schedule UTP in 2011 and not to file in 2012, the large temporary BTD for interest income and interest expense in both years are not associated with the decision to file Schedule UTP in 2011 and not to file in 2012. In both 2011 and 2012 the large temporary BTD on interest income approximately offset the large temporary BTD on interest expense with the result that the impact on total BTD for 2011 filers and 2012 nonfilers was small.

2. IRC Section 482 Citers

Schedule UTP filers are required to list one to three applicable IRC sections for each listed UTP concise description. Schedule UTP, Part I, tabulates the IRC information for each current year UTP concise description listed in Schedule UTP, Part III. SOI tabulates the first 10 rows of data on Schedule UTP, Part I. The five IRC sections appearing most frequently in the SOI file for Tax Years 2011 and 2012 are: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized costs). In 2012, seventy-two percent of such filers cited at least one of the five IRC sections. We categorize Schedule UTP filers as citing or not citing a specific IRC section. In this part of our study we focus on data describing Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 482.⁴⁴

Tables 4A and 4B and Figures 4A and 4B present 2012 data describing characteristics of Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 482. Figures 4A and 4B include 2011 data for comparison. Figure 4C presents 2011 and 2012 average data.

Schedule UTP filers citing IRC section 482 are generally larger than noncitors when study is restricted to the 1,176 Schedule UTP filers in 2012 with SEC 10K/Public financial statements and \$100 million or more in assets. The mean asset size as reported on the Form 1120 Schedule L by such Schedule UTP filers is \$23,921 million for the 326 filers citing IRC section 482 and \$20,157 million for the 850 filers not citing IRC section 482 (see Table 4B). The mean worldwide financial net income of the citers (reported on Schedule M-3, Part I, line 4) is \$1,017 million compared to \$374 million for the noncitors. The mean foreign nonincludible income of citers (reported on Schedule M-3, Part I, line 5) is -\$1,410 million (shown as negative since the income is removed in calculating book income) compared to -\$125 million for nonfilers. After the required Schedule M-3, Part I, adjustments, the mean book income for citers (reported on Schedule M-3, Part I, line 11) is \$631 million compared to \$410 million for the noncitors. Adding back U.S. Federal tax expense, mean pretax book income is \$804 million for the citers compared to \$498 million for the noncitors. Citers introduce mean BTD of -\$11 million to adjust pretax book to mean tax net income (reported on Schedule M-3, Part II, line 30, column D, and on Form 1120 page 1, line 28) of \$793 million compared to noncitors that introduce mean BTD of -\$79 million to adjust pretax book to mean tax net income of \$419 million.⁴⁵

We scale our initial aggregate dollar data by the adjusted-total-income measure we develop for our Mini M-3 analysis to make data for filers that cite or do not cite a specific IRC section more comparable. We used the adjusted-total-income book amount as a common-size scaling factor and compare percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers that cite or do not cite a specific IRC section, total pretax income BTD is expressed as a percentage of total pretax book income (see Tables 4A and 4B).

Table 4B shows that Schedule UTP filers in 2012 having SEC 10K/Public financial statements and \$100 million or more in assets, and citing IRC section 482 report:

- more worldwide income scaled as a percentage of adjusted total income than similar filers not citing IRC section 482 (24.03 percent compared to 13.76 percent);
- more scaled foreign nonincludible income (-33.31 percent versus -4.59 percent);
- *less* scaled book income (14.91 percent versus 15.09 percent);
- more scaled pretax book income (18.98 percent versus 18.30 percent);
- more scaled tax net income (18.73 percent versus 15.41 percent); but
- *less* negative scaled BTD (-0.25 percent versus -2.89 percent).

⁴⁴ We focus on Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more but with Audited (but not SEC 10K/Public) or Unaudited financial statements.

⁴⁵ Negative total BTD adjustments added to pretax book income result in tax net income that is lower than pretax book income.

TABLE 4A. 2012 U.S. Corporation 1120 Schedule M-3: Mini M-3—IRC 482 Cited: SEC 10K/Public: Assets \$100 Million or More

	\$ Million				Percentage of Adjusted Total Income				Percentage of Pretax Book					
	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Column B Temporary	Column C Permanent	Column D Tax	Column B Temporary	Column C Permanent	Column D Tax
	Total	Difference	Total	Difference	Total	Difference	Total	Difference	Total	Difference	Total	Difference	Total	Difference
UTP Filers Citing IRC 482														
Adj. Other income—no difference	2,504,196	0	0	2,504,196	0	181.42%	0.00%	181.42%	0.00%	0.00%	181.42%	0.00%	0.00%	0.00%
Adjusted COGS	-1,618,979	-841	-420	-1,620,239	-1,261	-117.29%	-0.06%	-117.38%	-0.09%	-0.32%	-117.38%	-0.32%	-0.16%	-0.48%
Adjusted gross profit	885,217	-841	-420	883,957	-1,261	64.13%	-0.06%	64.04%	-0.09%	-0.32%	64.04%	-0.32%	-0.16%	-0.48%
Specified income	316,896	22,232	-2,769	336,358	19,464	22.96%	1.61%	24.37%	1.41%	8.48%	24.37%	1.41%	-1.06%	7.43%
Adj. Other income—difference	178,192	9,659	-14,290	173,560	-4,632	12.91%	0.70%	12.57%	-0.34%	3.69%	12.57%	-0.34%	-5.45%	-1.77%
Adjusted total income	1,380,305	31,050	-17,479	1,393,875	13,571	100.00%	2.25%	100.98%	0.98%	11.85%	100.98%	0.98%	-6.67%	5.18%
Specified expense/deduction	-408,004	-11,419	70	-419,352	-11,349	-29.56%	-0.83%	-30.38%	-0.82%	-4.36%	-30.38%	-0.82%	0.03%	-4.33%
Other exp./ded.—difference	-366,729	-5,317	-338	-372,383	-5,655	-26.57%	-0.39%	-26.98%	-0.41%	-2.03%	-26.98%	-0.41%	-0.13%	-2.16%
Other exp./ded.—no difference	-343,541	0	0	-343,541	0	-24.89%	0.00%	-24.89%	0.00%	0.00%	-24.89%	0.00%	0.00%	0.00%
Pretax net income	262,031	14,314	-17,747	258,599	-3,433	18.98%	1.04%	18.73%	-0.25%	5.46%	18.73%	-0.25%	-6.77%	-1.31%
UTP Filers Not Citing IRC 482														
Adj. Other income—no difference	4,111,415	0	0	4,111,415	0	177.82%	0.00%	177.82%	0.00%	0.00%	177.82%	0.00%	0.00%	0.00%
Adjusted COGS	-2,452,101	-6,165	219	-2,458,047	-5,946	-106.05%	-0.27%	-106.31%	-0.26%	-1.46%	-106.31%	-0.26%	0.05%	-1.41%
Adjusted gross profit	1,659,314	-6,165	219	1,653,368	-5,946	71.76%	-0.27%	71.51%	-0.26%	-1.46%	71.51%	-0.26%	0.05%	-1.41%
Specified income	682,415	7,674	-29,493	660,467	-21,819	29.51%	0.33%	28.56%	-0.94%	1.81%	28.56%	-0.94%	-6.97%	-5.16%
Adj. Other income—difference	-29,574	14,451	-2,442	-17,565	12,009	-1.28%	0.63%	-0.76%	0.52%	3.41%	-0.76%	0.52%	-0.58%	2.84%
Adjusted total income	2,312,155	15,960	-31,716	2,296,270	-15,756	100.00%	0.69%	99.31%	-0.68%	3.77%	99.31%	-0.68%	-7.49%	-3.72%
Specified expense/deduction	-579,000	-24,341	8,175	-595,038	-16,166	-25.04%	-1.05%	-25.74%	-0.70%	-5.75%	-25.74%	-0.70%	1.93%	-3.82%
Other exp./ded.—difference	-723,822	-35,459	457	-758,824	-35,002	-31.31%	-1.53%	-32.82%	-1.51%	-8.38%	-32.82%	-1.51%	0.11%	-8.27%
Other exp./ded.—no difference	-586,132	0	0	-586,132	0	-25.35%	0.00%	-25.35%	0.00%	0.00%	-25.35%	0.00%	0.00%	0.00%
Pretax net income	423,201	-43,840	-23,084	356,276	-66,924	18.30%	-1.90%	15.41%	-2.89%	-10.36%	15.41%	-2.89%	-5.45%	-15.81%

If the BTD is scaled by pretax book income, the filers citing IRC section 482 reduce pretax book income by -1.31 percent to determine tax net income compared to a reduction of -15.81 percent by noncitors (see Table 4A, last row, last column, in each panel).

TABLE 4B. 2012 U.S. Corporation 1120 Schedule M-3: M-3, Part I—IRC 482 Cited: SEC 10K/Public: Assets \$100 Million or More

	Mean Amount Reported (\$ Million)		
	IRC 482 Cited	IRC 482 Not Cited	Difference
Worldwide financial net income	1,017	374	643
(Foreign nonincludible income)	-1,410	-125	-1,286
(U.S. nonincludible income)	-16	-38	22
Other includible income	-1	2	-3
Adjustments to eliminations	836	87	749
Other adjustments	183	110	73
Book net income includible corp.	631	410	221
Pretax net income—book	804	498	306
Tax net income	793	419	374
Total BTD difference	-11	-79	68
Total assets—Schedule L	23,921	20,157	3,764
Assets—Financial statements	24,143	17,000	7,143
Assets—Foreign nonincludible	14,961	4,737	10,223
Assets—U.S. nonincludible	757	612	145
Assets—Other includible	137	110	27
	Percentage of Adjusted Total Income Book		
	IRC 482 Cited	IRC 482 Not Cited	Difference
Worldwide financial net income	24.03%	13.76%	10.26%
(Foreign nonincludible income)	-33.31%	-4.59%	-28.73%
(U.S. nonincludible income)	-0.37%	-1.39%	1.02%
Other includible income	-0.03%	0.07%	-0.10%
Adjustments to eliminations	19.74%	3.18%	16.56%
Other adjustments	4.33%	4.06%	0.28%
Book net income includible corp.	14.91%	15.09%	-0.18%
Pretax net income—book	18.98%	18.30%	0.68%
Tax net income	18.73%	15.41%	3.33%
Total BTD difference	-0.25%	-2.89%	2.65%
	Distribution of Returns		
	IRC 482 Cited	IRC 482 Not Cited	Total
Number of returns	326	850	1,176
Share of returns	27.7%	72.3%	100.0%

Figure 4A compares Schedule M-3 data as percentages of adjusted total income for 2011 and 2012 Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 482, for seven key items: worldwide financial income (Part I, line 4); foreign nonincludible income (Part I, line 5); adjustments to eliminations (Part I, line 8); other adjustments (Part I, line 10); book income (Part I, line 11); pretax book income; and tax net income. With the exception of book income for 2012, for both years, filers citing IRC section 482 report larger scaled amounts for the seven items than noncitors.

Figure 4B uses pretax book income for scaling 2011 and 2012 BTD, temporary, permanent, and total, for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 482. In 2011, citers have less negative temporary BTD than noncitors, and have positive permanent BTD compared to negative permanent BTD for noncitors. In 2012, citers have positive temporary BTD compared to negative temporary BTD for noncitors, and have more negative permanent BTD than noncitors. In aggregate, the citers report approximately offsetting temporary and permanent BTD. Citers have less negative total BTD than noncitors after the offsetting effects of the temporary BTD and permanent BTD for the citers. Stated differently, the otherwise similar noncitors use negative total BTD to reduce tax net income more than the citers.

FIGURE 4A. 2011–2012 U.S. Corporation Key M-3 Data as Percentage of Adjusted Total Book Income for SEC/10K Public Financial Statements by IRC Section 482

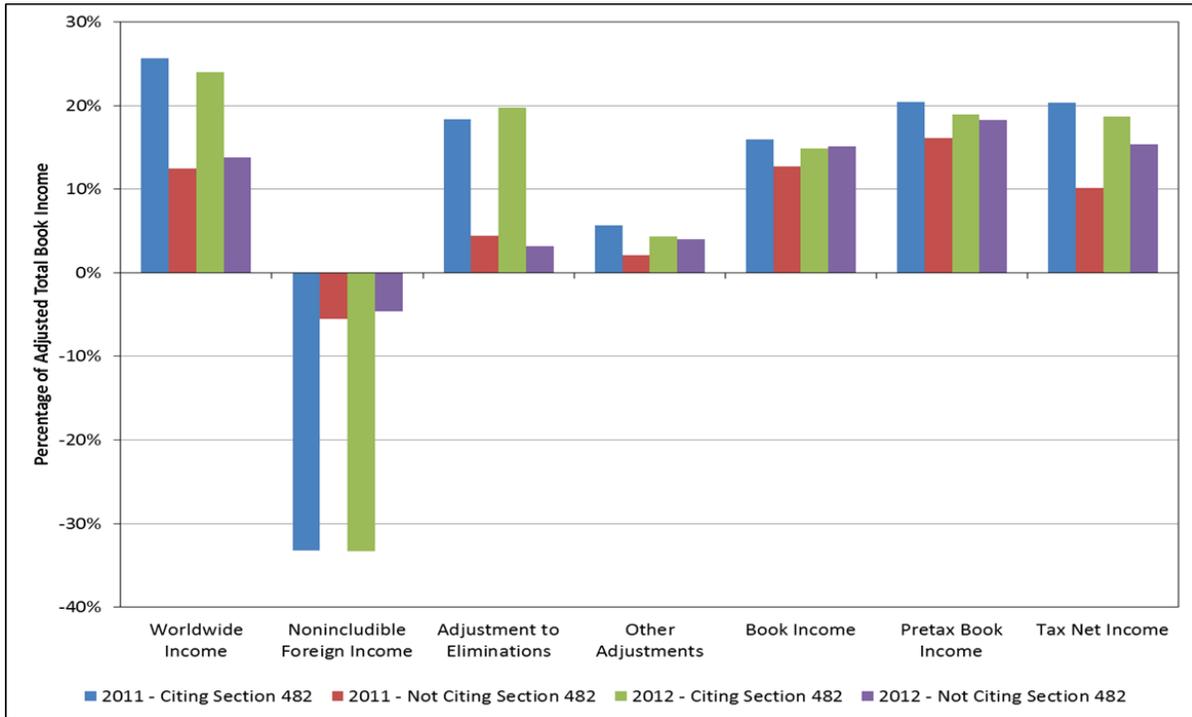


FIGURE 4B. 2011–2012 U.S. Corporation M-3: BTD as Percentage of Pretax Book Income for SEC/10K Public Financial Statements by IRC Section 482

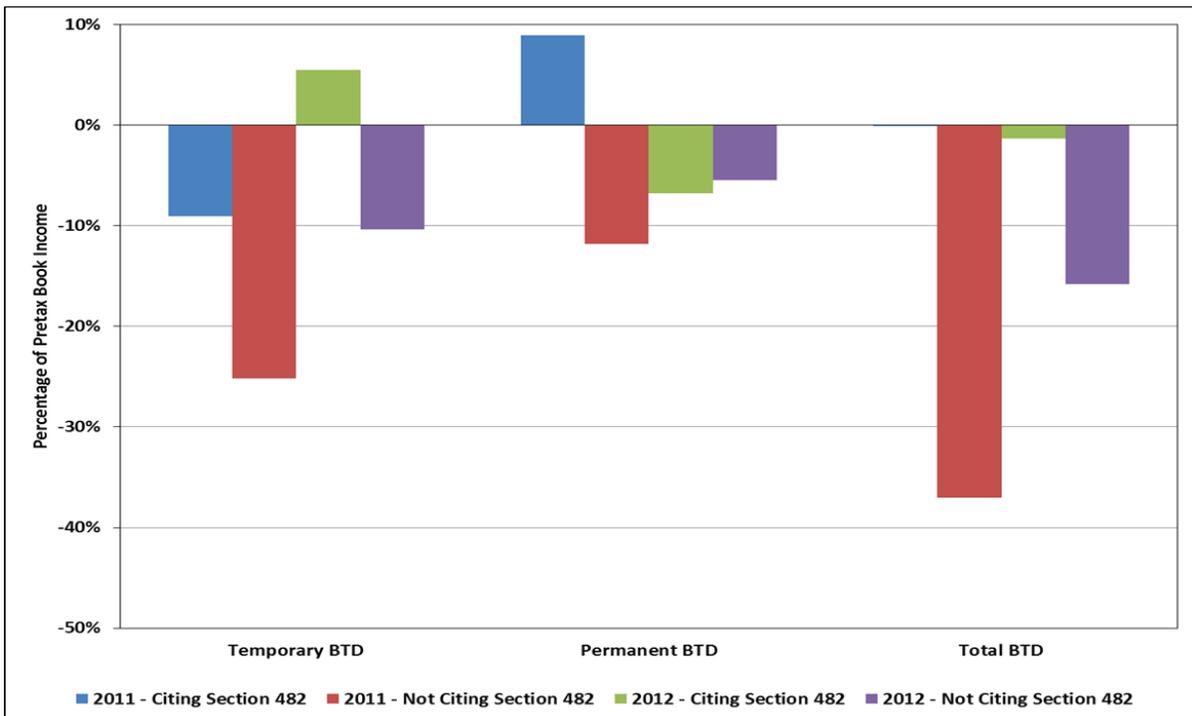
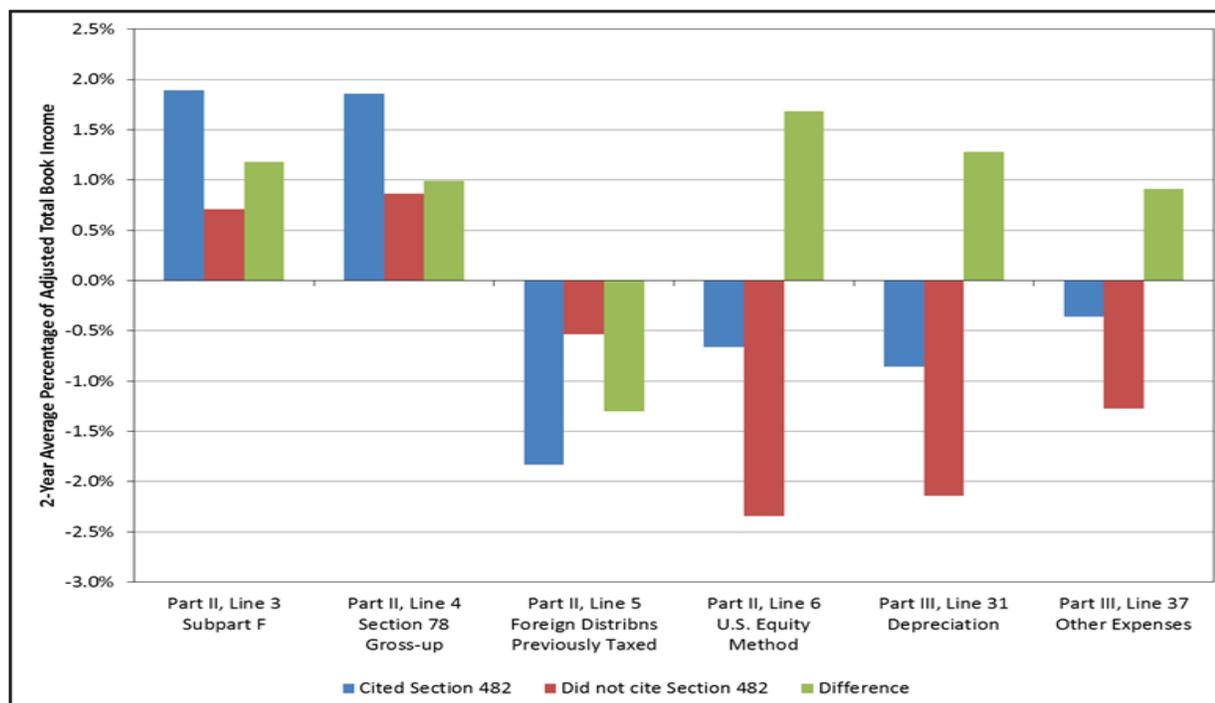


FIGURE 4C. 2011–2012 U.S. Corporation M-3: 2-Year Average Total BTD as Percentage of Adjusted Total Book Income for Top 6 Lines—UTP Filers Citing/Not Citing IRC 482 (SEC/10K Public Financial Statements)



In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

Figure 4C reports total BTD for six specific Schedule M-3 lines for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 482.⁴⁶ The total BTD for each of the six lines are expressed as a percentage of adjusted total income and averaged for 2011 and 2012. The lines are the top (that is, the most extreme) six lines on Schedule M-3, Parts II and III, in terms of the absolute value of the difference in 2-year-average scaled BTD between citers and noncitters.⁴⁷ The percentages for each of these lines are provided below.

Schedule M-3 Lines Featured in Figure 4C		Citers	Nonciters	Difference
Part II, line 3:	Inclusion in tax income of subpart F foreign income	1.89%	0.71%	1.18%
Part II, line 4:	Inclusion in tax income of Section 78 gross-up	1.86%	0.86%	0.99%
Part II, line 5:	Exclusion in tax income of previously taxed foreign distributions	-1.83%	-0.54%	-1.30%
Part II, line 6:	Exclusion from tax income of U.S. equity method income	-0.66%	-2.34%	1.68%
Part III, line 31:	Adjustment to depreciation expense/deduction in tax income	-0.86%	-2.14%	1.28%
Part III, line 37:	Adjustment to other expense/deduction with difference in tax income	-0.36%	-1.28%	0.91%

⁴⁶ The six lines are selected from and represent approximately 10 percent of the 68 specified and other-with-difference lines on Schedule M-3, Parts II and III. See the discussion and footnotes for Figure 3C.

⁴⁷ For the current study, we exclude interest income (Part II, line 13) and interest expense (Part III, line 8) as key lines because of a data anomaly in the 2011 and 2012 data. See the discussion and footnotes for Figure 3C.

3. IRC Section 41 Citers

Schedule UTP filers are required to list one to three applicable IRC sections for each listed UTP concise description. Schedule UTP, Part I, tabulates the IRC information for each current year UTP concise description listed in Schedule UTP, Part III. SOI tabulates the first 10 rows of data on Schedule UTP, Part I. The five IRC sections appearing most frequently in the SOI file for Tax Years 2011 and 2012 are: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized costs). In 2012, seventy-two percent of such filers cited at least one of the five IRC sections. We categorize Schedule UTP filers as citing or not citing a specific IRC section. In this part of our study we focus on data describing Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 41.⁴⁸

Tables 5A and 5B and Figures 5A and 5B present 2012 data describing characteristics of Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 41. Figures 5A and 5B include 2011 data for comparison. Figure 5C presents 2011 and 2012 average data.

Schedule UTP filers citing IRC section 41 are generally smaller than noncitors when study is restricted to the 1,176 Schedule UTP filers in 2012 with SEC 10K/Public financial statements and \$100 million or more in assets. The mean asset size as reported on the Form 1120, Schedule L, by such Schedule UTP filers is \$14,237 million for the 506 filers citing IRC section 41 and \$26,459 million for the 670 filers not citing IRC section 41 (see Table 5B). The mean worldwide financial net income of the citers (reported on Schedule M-3, Part I, line 4) is \$513 million compared to \$583 million for the noncitors. The mean foreign nonincludible income of citers (reported on Schedule M-3, Part I, line 5) is -\$487 million (shown as negative since the income is removed in calculating book income) compared to -\$477 million for nonfilers. After the required Schedule M-3, Part I, adjustments, the mean book income for citers (reported on Schedule M-3, Part I, line 11) is \$420 million compared to \$511 million for the noncitors. Adding back U.S. Federal tax expense, mean pretax book income is \$508 million for the citers compared to \$639 million for the noncitors. Citers introduce mean BTD of -\$138 million to adjust pretax book to mean tax net income (reported on Schedule M-3, Part II, line 30, column D, and on Form 1120, page 1, line 28) of \$370 million compared to noncitors that introduce mean BTD of -\$1 million to adjust pretax book to mean tax net income of \$638 million.⁴⁹

We scale our initial aggregate dollar data by the adjusted-total-income measure we develop for our Mini M-3 analysis to make data for filers that cite or do not cite a specific IRC section more comparable. We used the adjusted-total-income book amount as a common-size scaling factor and compare percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers that cite or do not cite a specific IRC section, total pretax income BTD is expressed as a percentage of total pretax book income (see Tables 5A and 5B).

Table 5B shows that Schedule UTP filers in 2012 having SEC 10K/Public financial statements and \$100 million or more in assets, and citing IRC section 41 report:

- more worldwide income scaled as a percentage of adjusted total income than similar filers not citing IRC section 41 (21.10 percent compared to 15.85 percent);
- more scaled foreign nonincludible income (-20.05 percent versus -12.97 percent);
- more scaled book income (17.29 percent versus 13.89 percent);
- more scaled pretax book income (20.91 percent versus 17.38 percent);
- less scaled tax net income (15.25 percent versus 17.35 percent); and
- more negative scaled BTD (-5.66 percent versus -0.03 percent).

⁴⁸ We focus on Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more but with Audited (but not SEC 10K/Public) or Unaudited financial statements.

⁴⁹ Negative total BTD adjustments added to pretax book income result in tax net income that is lower than pretax book income.

TABLE 5A. 2012 U.S. Corporation 1120 Schedule M-3: Mini M-3—IRC 41 Cited: SEC 10K/Public: Assets \$100 Million or More

	\$ Million						Percentage of Adjusted Total Income						Percentage of Pretax Book		
	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Total Difference	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Total Difference	Column B Temporary	Column C Permanent	Total Difference		
UTP Filers Citing IRC 41															
Adj. Other income—no difference	2,186,076	0	0	2,186,076	0	177.82%	0.00%	0.00%	177.82%	0.00%	0.00%	0.00%	0.00%		
Adjusted COGS	-1,320,017	-1,233	35	-1,321,213	-1,198	-107.37%	-0.10%	0.00%	-107.47%	-0.10%	-0.48%	0.01%	-0.47%		
Adjusted gross profit	866,059	-1,233	35	864,863	-1,198	70.45%	-0.10%	0.00%	70.35%	-0.10%	-0.48%	0.01%	-0.47%		
Specified income	610,166	394	-41,824	568,727	-41,430	49.63%	0.03%	-3.40%	46.26%	-3.37%	0.15%	-16.27%	-16.12%		
Adj. Other income—difference	-246,846	-2,285	-8,166	-257,297	-10,451	-20.08%	-0.19%	-0.66%	-20.93%	-0.85%	-0.89%	-3.18%	-4.07%		
Adjusted total income	1,229,379	-3,124	-49,955	1,176,293	-53,079	100.00%	-0.25%	-4.06%	95.68%	-4.32%	-1.22%	-19.43%	-20.65%		
Specified expense/deduction	-316,779	-8,578	472	-324,877	-8,106	-25.77%	-0.70%	0.04%	-26.43%	-0.66%	-3.34%	0.18%	-3.15%		
Other exp./ded.—difference	-396,530	-6,331	-2,117	-404,978	-8,448	-32.25%	-0.51%	-0.17%	-32.94%	-0.69%	-2.46%	-0.82%	-3.29%		
Other exp./ded.—no difference	-259,004	0	0	-259,004	0	-21.07%	0.00%	0.00%	-21.07%	0.00%	0.00%	0.00%	0.00%		
Pretax net income	257,066	-18,033	-51,600	187,434	-69,633	20.91%	-1.47%	-4.20%	15.25%	-5.66%	-7.01%	-20.07%	-27.09%		
UTP Filers Not Citing IRC 41															
Adj. Other income—no difference	4,429,536	0	0	4,429,536	0	179.84%	0.00%	0.00%	179.84%	0.00%	0.00%	0.00%	0.00%		
Adjusted COGS	-2,751,063	-5,773	-237	-2,757,073	-6,010	-111.69%	-0.23%	-0.01%	-111.94%	-0.24%	-1.35%	-0.06%	-1.40%		
Adjusted gross profit	1,678,473	-5,773	-237	1,672,463	-6,010	68.15%	-0.23%	-0.01%	67.90%	-0.24%	-1.35%	-0.06%	-1.40%		
Specified income	389,146	29,513	9,562	428,099	39,075	15.80%	1.20%	0.39%	17.38%	1.59%	6.89%	2.23%	9.13%		
Adj. Other income—difference	395,461	26,395	-8,566	413,290	17,829	16.06%	1.07%	-0.35%	16.78%	0.72%	6.16%	-2.00%	4.16%		
Adjusted total income	2,463,080	50,135	759	2,513,852	50,894	100.00%	2.04%	0.03%	102.06%	2.07%	11.71%	0.18%	11.89%		
Specified expense/deduction	-670,225	-27,182	7,773	-689,513	-19,409	-27.21%	-1.10%	0.32%	-27.99%	-0.79%	-6.35%	1.82%	-4.53%		
Other exp./ded.—difference	-694,020	-34,445	2,235	-726,230	-32,209	-28.18%	-1.40%	0.09%	-29.48%	-1.31%	-8.04%	0.52%	-7.52%		
Other exp./ded.—no difference	-670,669	0	0	-670,669	0	-27.23%	0.00%	0.00%	-27.23%	0.00%	0.00%	0.00%	0.00%		
Pretax net income	428,166	-11,492	10,767	427,440	-724	17.38%	-0.47%	0.44%	17.35%	-0.03%	-2.68%	2.51%	-0.17%		

If the BTD is scaled by pretax book income, the filers citing IRC section 41 reduce pretax book income by -27.09 percent to determine tax net income compared to a reduction of -0.17 percent by noncitors (see Table 5A, last row, last column, in each panel). In short, those SEC 10K/Public Schedule UTP filers that cite IRC section 41 are both more profitable and declare less tax net income than those that do not cite IRC section 41.

TABLE 5B. 2012 U.S. Corporation 1120 Schedule M-3: M-3, Part I—IRC 41 Cited: SEC 10K/ Public: Assets \$100 Million or More

	Mean Amount Reported (\$ Million)		
	IRC 41 Cited	IRC 41 Not Cited	Difference
Worldwide financial net income	513	583	-70
(Foreign nonincludible income)	-487	-477	-10
(U.S. nonincludible income)	-9	-49	39
Other includible income	2	0	2
Adjustments to eliminations	366	240	125
Other adjustments	36	202	-166
Book net income includible corp.	420	511	-91
Pretax net income—book	508	639	-131
Tax net income	370	638	-268
Total BTD difference	-138	-1	-137
Total assets—Schedule L	14,237	26,459	-12,223
Assets—Financial statements	13,715	22,957	-9,242
Assets—Foreign nonincludible	7,053	7,963	-910
Assets—U.S. nonincludible	445	809	-365
Assets—Other includible	12	197	-186
	Percentage of Adjusted Total Income Book		
	IRC 41 Cited	IRC 41 Not Cited	Difference
Worldwide financial net income	21.10%	15.85%	5.25%
(Foreign nonincludible income)	-20.05%	-12.97%	-7.08%
(U.S. nonincludible income)	-0.37%	-1.32%	0.95%
Other includible income	0.09%	0.00%	0.09%
Adjustments to eliminations	15.05%	6.54%	8.51%
Other adjustments	1.47%	5.50%	-4.03%
Book net income includible corp.	17.29%	13.89%	3.39%
Pretax net income—book	20.91%	17.38%	3.53%
Tax net income	15.25%	17.35%	-2.11%
Total BTD difference	-5.66%	-0.03%	-5.63%
	Distribution of Returns		
	IRC 41 Cited	IRC 41 Not Cited	Total
Number of returns	506	670	1,176
Share of returns	43.0%	57.0%	100.0%

Figure 5A compares Schedule M-3 data as percentages of adjusted total income for 2011 and 2012 Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 41, for seven key items: worldwide financial income (Part I, line 4); foreign nonincludible income (Part I, line 5); adjustments to eliminations (Part I, line 8); other adjustments (Part I, line 10); book income (Part I, line 11); pretax book income; and tax net income. With the exception of other adjustments for 2011 and 2012 and tax net income for 2012, for both years, filers citing IRC section 41 report larger scaled amounts for the seven items than noncitors.

Figure 5B uses pretax book income for scaling 2011 and 2012 BTD, temporary, permanent, and total, for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 41. In 2011, citers have less negative temporary BTD than noncitors, and have negative permanent BTD compared to positive permanent BTD for noncitors. In 2012, citers have more negative

FIGURE 5A. 2011–2012 U.S. Corporation Key M-3 Data as Percentage of Adjusted Total Book Income for SEC/10K Public Financial Statements by IRC Section 41

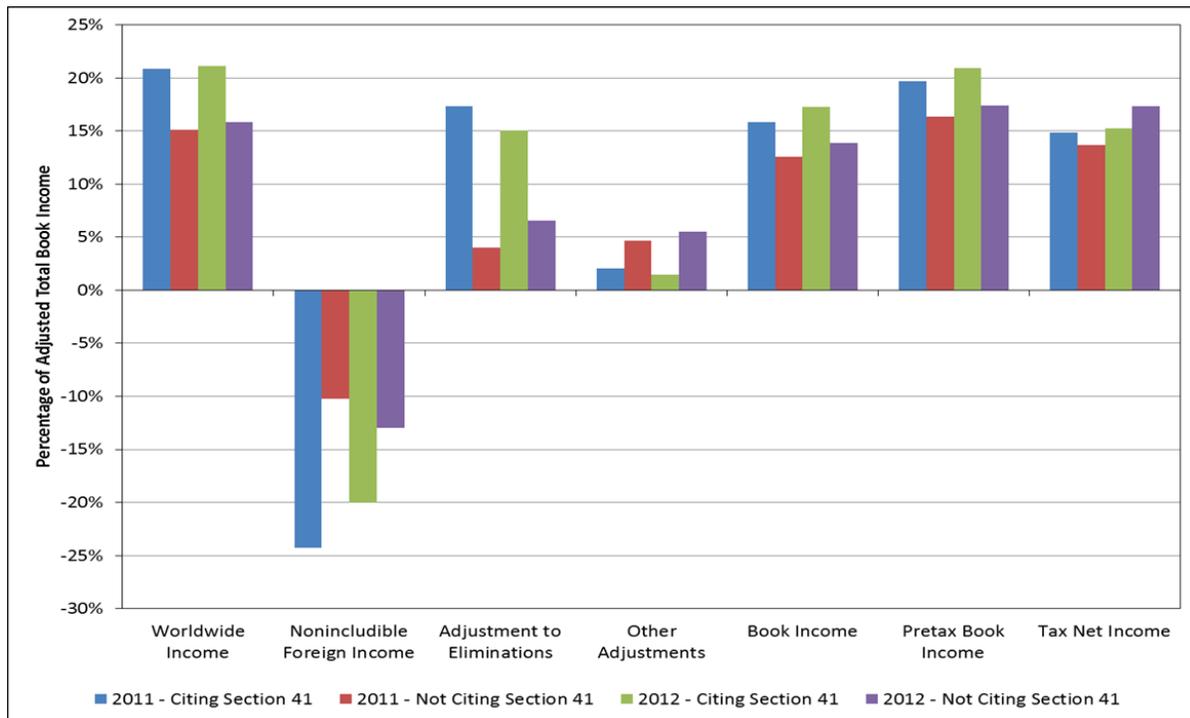


FIGURE 5B. 2011–2012 U.S. Corporation M-3: BTD as Percentage of Pretax Book Income for SEC/10K Public Financial Statements by IRC Section 41

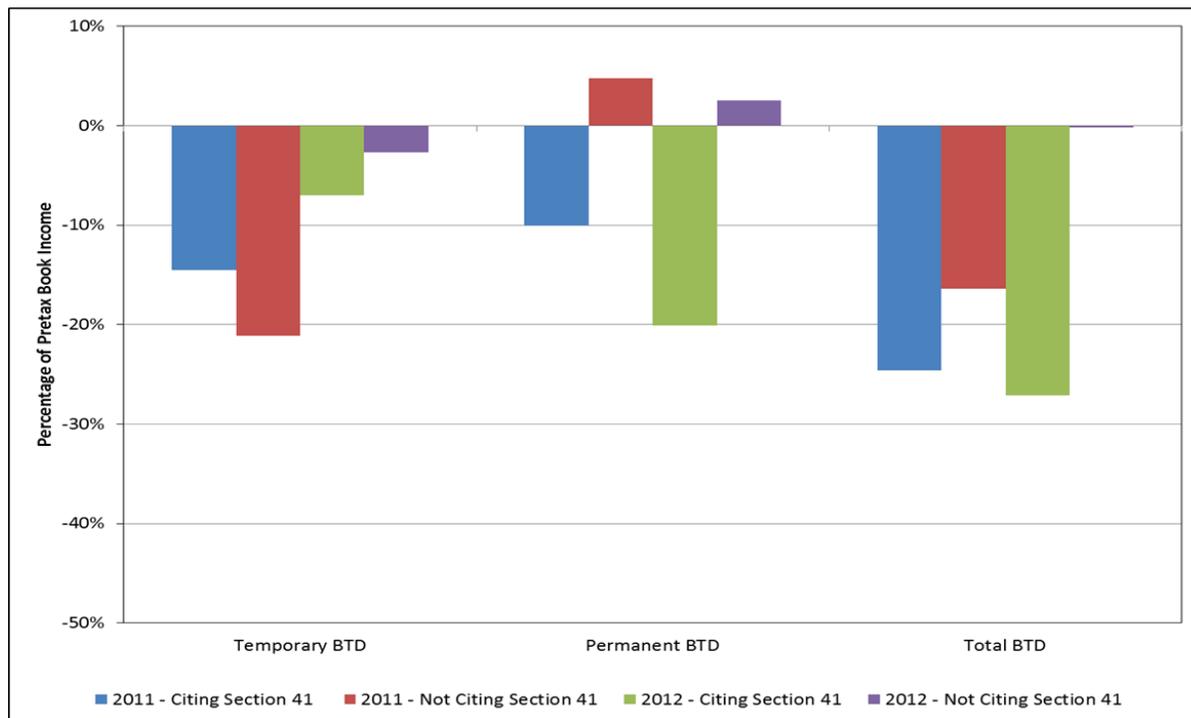
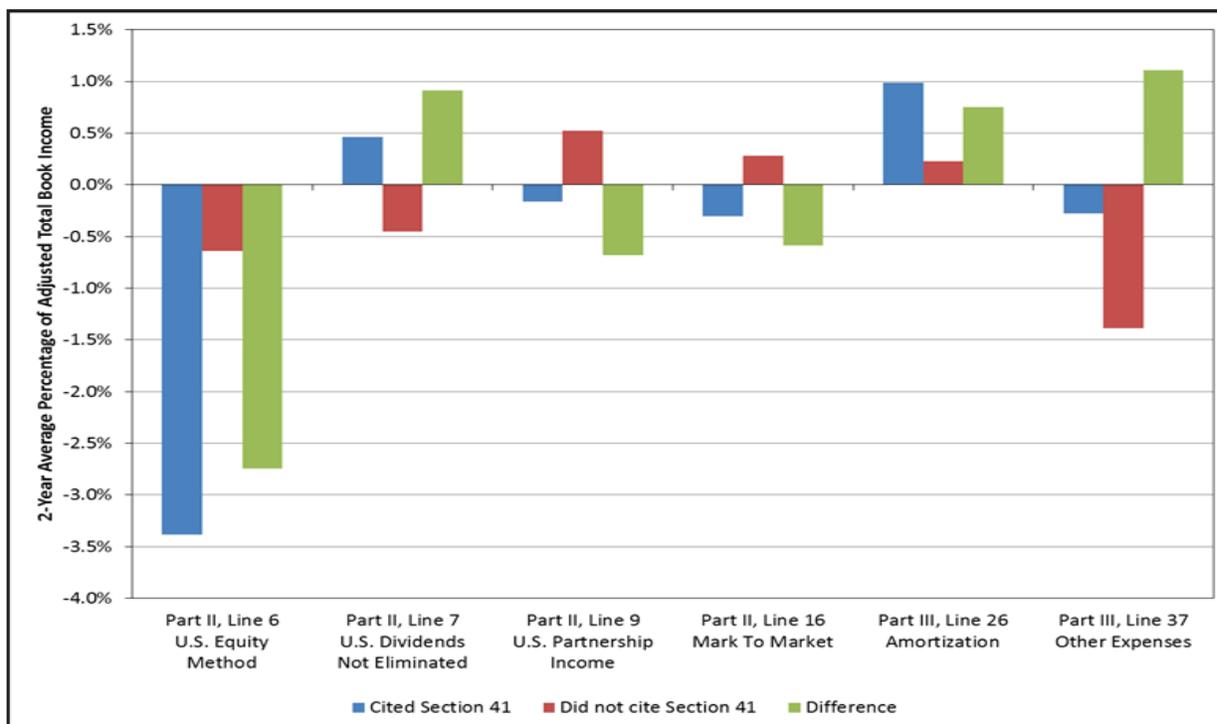


FIGURE 5C. 2011–2012 U.S. Corporation M-3: 2-Year Average Total BTD as Percentage of Adjusted Total Book Income for Top 6 Lines—UTP Filers Citing/Not Citing IRC Section 41 (SEC/10K Public Financial Statements)



temporary BTB than nonciters, and have negative permanent BTB compared to positive permanent BTB for nonciters. In aggregate, citers of IRC section 41 have more negative total BTB than nonciters in both 2011 and 2012. Stated differently, the otherwise similar nonciters use negative total BTB to reduce tax net income less than the citers.

In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

Figure 5C reports total BTB for six specific Schedule M-3 lines for Schedule UTP filers, with SEC 10K/ Public financial statements and \$100 million or more in assets, citing or not citing IRC section 41.⁵⁰ The total BTB for each of the six lines are expressed as a percentage of adjusted total income and averaged for 2011 and 2012. The lines are the top (that is, the most extreme) six lines on Schedule M-3, Parts II and III, in terms of the absolute value of the difference in 2-year-average scaled BTB between citers and nonciters.⁵¹ The percentages for each of these lines are provided below.

⁵⁰ The six lines are selected from and represent approximately 10 percent of the 68 specified and other-with-difference lines on Schedule M-3, Parts II and III. See the discussion and footnotes for Figure 3C.

⁵¹ For the current study, we exclude interest income (Part II, line 13) and interest expense (Part III, line 8) as key lines because of a data anomaly in the 2011 and 2012 data. See the discussion and footnotes for Figure 3C.

Schedule M-3 Lines Featured in Figure 5C		Citers	Nonciters	Difference
Part II, line 6:	Exclusion from tax income of U.S. equity method income	-3.39%	-0.64%	-2.75%
Part II, line 7:	Adjustments to U.S. dividends, not eliminated in consolidation, in tax income	0.46%	-0.45%	0.92%
Part II, line 9:	Adjustments to U.S. partnership income to include all Schedule K-1 income in tax income	-0.16%	0.52%	-0.68%
Part II, line 16:	Adjustments for mark-to-market in tax income	-0.30%	0.28%	-0.59%
Part III, line 26:	Adjustment to amortization/impairment of goodwill expense/deduction in tax income	0.98%	0.23%	0.75%
Part III, line 37:	Adjustment to other expense/deduction with difference in tax income	-0.28%	-1.38%	1.11%

Schedule UTP filers citing IRC section 162 are generally larger than nonciters when study is restricted to the 1,176 Schedule UTP filers in 2012 with SEC 10K/Public financial statements and \$100 million or more in assets. The mean asset size as reported on the Form 1120, Schedule L, by such Schedule UTP filers is \$52,298 million for the 164 filers citing IRC section 162 and \$16,160 million for the 1,012 filers not citing IRC section 162 (see Table 6B). The mean worldwide financial net income of the citers (reported on Schedule M-3, Part I, line 4) is \$948 million compared to \$488 million for the nonciters. The mean foreign nonincludible income of citers (reported on Schedule M-3, Part I, line 5) is -\$380 million (shown as negative since the income is removed in calculating book income) compared to -\$497 million for nonfilers. After the required Schedule M-3, Part I, adjustments, the mean book income for citers (reported on Schedule M-3, Part I, line 11) is \$1,004 million compared to \$385 million for the nonciters. Adding back U.S. Federal tax expense, mean pretax book income is \$1,244 million for the citers compared to \$475 million for the nonciters. Citers introduce mean BTD of -\$90 million to adjust pretax book to mean tax net income (reported on Schedule M-3, Part II, line 30, column D, and on Form 1120, page 1, line 28) of \$1,154 million compared to nonciters that introduce mean BTD of -\$55 million to adjust pretax book to mean tax net income of \$421 million.⁵²

We scale our initial aggregate dollar data by the adjusted-total-income measure we develop for our Mini M-3 analysis to make data for filers that cite or do not cite a specific IRC section more comparable. We used the adjusted-total-income book amount as a common-size scaling factor and compare percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers that cite or do not cite a specific IRC section, total pretax income BTD is expressed as a percentage of total pretax book income (see Tables 6A and 6B).

4. IRC Section 162 Citers

Schedule UTP filers are required to list one to three applicable IRC sections for each listed UTP concise description. Schedule UTP, Part I, tabulates the IRC information for each current year UTP concise description listed in Schedule UTP, Part III. SOI tabulates the first 10 rows of data on Schedule UTP, Part I. The five IRC sections appearing most frequently in the SOI file for Tax Years 2011 and 2012 are: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized costs). In 2012, seventy-two percent of such filers cited at least one of the five IRC sections. We categorize Schedule UTP filers as citing or not citing a specific IRC section. In this part of our study we focus on data describing Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 162.⁵³

Tables 6A and 6B and Figures 6A and 6B present 2012 data describing characteristics of Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 162. Figures 6A and 6B include 2011 data for comparison. Figure 6C presents 2011 and 2012 average data.

⁵² We focus on Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more but with Audited (but not SEC 10K/Public) or Unaudited financial statements.

⁵³ Negative total BTB adjustments added to pretax book income result in tax net income that is lower than pretax book income.

TABLE 6A. 2012 U.S. Corporation 1120 Schedule M-3: Mini M-3—IRC 162 Cited: SEC 10K/Public: Assets \$100 Million or More

	\$ Million					Percentage Adjusted Total Income					Percentage Pretax Book		
	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Total Difference	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Total Difference	Column B Temporary	Column C Permanent	Total Difference
UTP Filers Citing IRC 162													
Adj. Other income—no difference	1,631,130	0	0	1,631,130	0	145.56%	0.00%	0.00%	145.56%	0.00%	0.00%	0.00%	0.00%
Adjusted COGS	-843,911	-2,620	234	-846,297	-2,386	-75.31%	-0.23%	0.02%	-75.52%	-0.21%	-1.28%	0.11%	-1.17%
Adjusted gross profit	787,219	-2,620	234	784,833	-2,386	70.25%	-0.23%	0.02%	70.04%	-0.21%	-1.28%	0.11%	-1.17%
Specified income	267,582	693	5,490	273,644	6,183	23.88%	0.06%	0.49%	24.42%	0.55%	0.34%	2.69%	3.03%
Adj. Other income—difference	65,758	16,679	-1,571	80,866	15,108	5.87%	1.49%	-0.14%	7.22%	1.35%	8.17%	-0.77%	7.40%
Adjusted total income	1,120,559	14,752	4,153	1,139,343	18,905	100.00%	1.32%	0.37%	101.68%	1.69%	7.23%	2.03%	9.26%
Specified expense/deduction	-281,815	-11,320	-485	-293,499	-11,805	-25.15%	-1.01%	-0.04%	-26.19%	-1.05%	-5.55%	-0.24%	-5.78%
Other exp./ded.—difference	-330,478	-21,829	-73	-352,379	-21,902	-29.49%	-1.95%	-0.01%	-31.45%	-1.95%	-10.70%	-0.04%	-10.73%
Other exp./ded.—no difference	-304,187	0	0	-304,187	0	-27.15%	0.00%	0.00%	-27.15%	0.00%	0.00%	0.00%	0.00%
Pretax net income	204,079	-18,397	3,595	189,278	-14,802	18.21%	-1.64%	0.32%	16.89%	-1.32%	-9.01%	1.76%	-7.25%
UTP Filers Not Citing IRC 162													
Adj. Other income—no difference	4,984,482	0	0	4,984,482	0	193.81%	0.00%	0.00%	193.81%	0.00%	0.00%	0.00%	0.00%
Adjusted COGS	-3,227,169	-4,386	-436	-3,231,989	-4,821	-125.48%	-0.17%	-0.02%	-125.67%	-0.19%	-0.91%	-0.09%	-1.00%
Adjusted gross profit	1,757,313	-4,386	-436	1,752,493	-4,821	68.33%	-0.17%	-0.02%	68.14%	-0.19%	-0.91%	-0.09%	-1.00%
Specified income	731,729	29,214	-37,752	723,181	-8,538	28.45%	1.14%	-1.47%	28.12%	-0.33%	6.07%	-7.85%	-1.77%
Adj. Other income—difference	82,858	7,431	-15,161	75,128	-7,730	3.22%	0.29%	-0.59%	2.92%	-0.30%	1.54%	-3.15%	-1.61%
Adjusted total income	2,571,900	32,259	-53,349	2,550,802	-21,089	100.00%	1.25%	-2.07%	99.18%	-0.82%	6.70%	-11.09%	-4.38%
Specified expense/deduction	-705,189	-24,440	8,730	-720,891	-15,710	-27.42%	-0.95%	0.34%	-28.03%	-0.61%	-5.08%	1.81%	-3.27%
Other exp./ded.—difference	-760,073	-18,947	191	-778,828	-18,756	-29.55%	-0.74%	0.01%	-30.28%	-0.73%	-3.94%	0.04%	-3.90%
Other exp./ded.—no difference	-625,486	0	0	-625,486	0	-24.32%	0.00%	0.00%	-24.32%	0.00%	0.00%	0.00%	0.00%
Pretax net income	481,152	-11,128	-44,428	425,597	-55,555	18.71%	-0.43%	-1.73%	16.55%	-2.16%	-2.31%	-9.23%	-11.55%

**TABLE 6B. 2012 U.S. Corporation 1120 Schedule M-3: M-3, Part I—IRC 162 Cited: SEC 10K/
Public: Assets \$100 Million or More**

	Mean Amount Reported (\$ Million)		
	IRC 162 Cited	IRC 162 Not Cited	Difference
Worldwide financial net income	948	488	460
(Foreign nonincludible income)	-380	-497	117
(US nonincludible income)	-54	-28	-26
Other includible income	-1	1	-2
Adjustments to eliminations	-75	354	-429
Other adjustments	565	60	505
Book net income includible corp.	1,004	385	619
Pretax net income—book	1,244	475	769
Tax net income	1,154	421	734
Total BTD difference	-90	-55	-35
Total assets—Schedule L	52,298	16,160	36,138
Assets—Financial statements	43,792	14,960	28,832
Assets—Foreign nonincludible	8,060	7,492	568
Assets—US nonincludible	1,033	591	442
Assets—Other includible	51	128	-78
	Percentage of Adjusted Total Income Book		
	IRC 162 Cited	IRC 162 Not Cited	Difference
Worldwide financial net income	13.88%	19.22%	-5.34%
(Foreign nonincludible income)	-5.57%	-19.58%	14.01%
(U.S. nonincludible income)	-0.79%	-1.10%	0.31%
Other includible income	-0.01%	0.05%	-0.05%
Adjustments to eliminations	-1.10%	13.94%	-15.04%
Other adjustments	8.27%	2.37%	5.91%
Book net income includible corp.	14.70%	15.16%	-0.46%
Pretax net income—book	18.21%	18.71%	-0.50%
Tax net income	16.89%	16.55%	0.34%
Total BTD difference	-1.32%	-2.16%	0.84%
	Distribution of Returns		
	IRC 162 Cited	IRC 162 Not Cited	Total
Number of returns	164	1,012	1,176
Share of returns	13.9%	86.1%	100.0%

Table 6B shows that Schedule UTP filers in 2012 having SEC 10K/Public financial statements and \$100 million or more in assets, and citing IRC section 162 report:

- *less* worldwide income scaled as a percentage of adjusted total income than similar filers not citing IRC section 162 (13.88 percent compared to 19.22 percent);
- less scaled foreign nonincludible income (-5.57 percent versus -19.58 percent);
- less scaled book income (14.70 percent versus 15.16 percent);
- less scaled pretax book income (18.21 percent versus 18.71 percent);
- *more* scaled tax net income (16.89 percent versus 16.55 percent); and
- less negative scaled BTD (-1.32 percent versus -2.16 percent).

If the BTD is scaled by pretax book income, the filers citing IRC section 162 reduce pretax book income by -7.25 percent to determine tax net income compared to a reduction of -11.55 percent by noncitters (see Table 6A, last row, last column, in each panel).

Figure 6A compares Schedule M-3 data as percentages of adjusted total income for 2011 and 2012 Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 162, for seven key items: worldwide financial income (Part I, line 4); foreign nonincludible income (Part I, line 5); adjustments to eliminations (Part I, line 8); other adjustments (Part I, line 10); book income (Part I, line 11); pretax book income; and tax net income. With the exception of worldwide income for 2011, other adjustments for 2012, book income for 2011, pretax book income for 2011, and tax net income for 2011 and 2012, for both years, filers citing IRC section 162 report smaller scaled amounts for the seven items than noncitters.

Figure 6B uses pretax book income for scaling 2011 and 2012 BTD, temporary, permanent, and total, for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 162. In 2011 and 2012, citers have more negative temporary BTD than noncitters, and have positive permanent BTD compared to negative permanent BTD for noncitters. In aggregate, citers of IRC section 162 have less negative total BTD than noncitters in both 2011 and 2012. Stated differently, the otherwise similar noncitters use negative total BTD to reduce tax net income more than the citers.

In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

FIGURE 6A. 2011–2012 U.S. Corporation Key M-3 Data as Percentage of Adjusted Total Book Income for SEC/10K Public Financial Statements by IRC Section 162

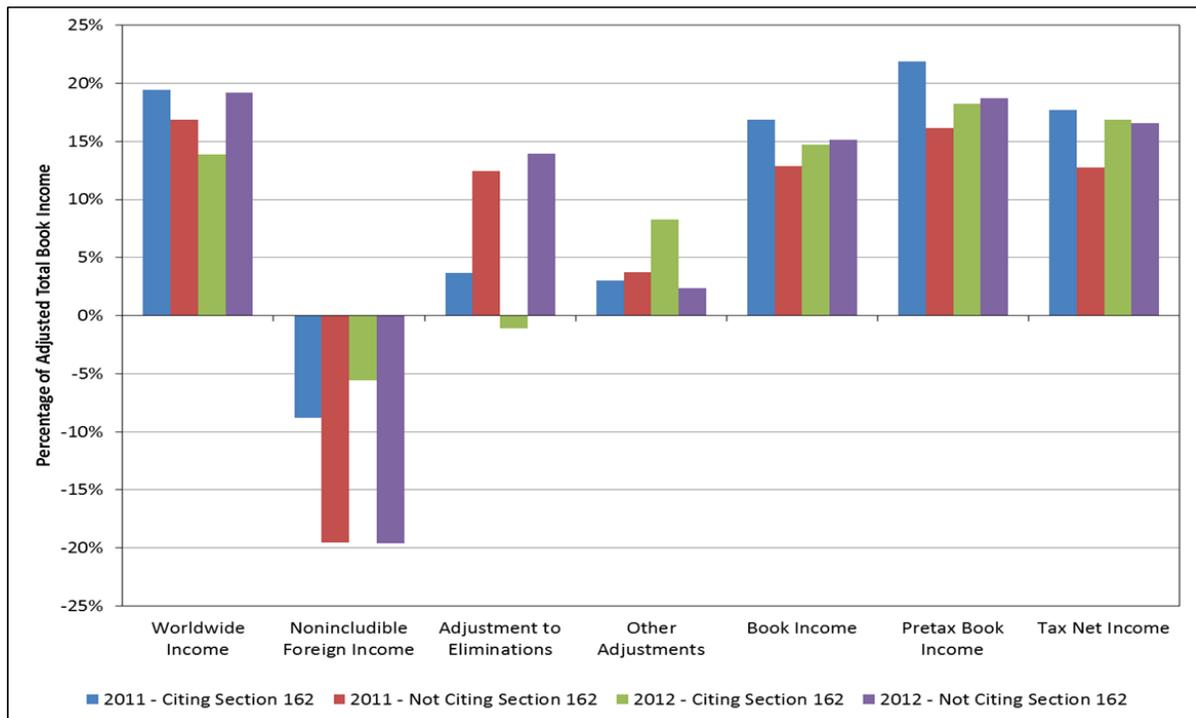


FIGURE 6B. 2011–2012 U.S. Corporation M-3: BTD as Percentage of Pretax Book Income for SEC/10K Public Financial Statements by IRC Section 162

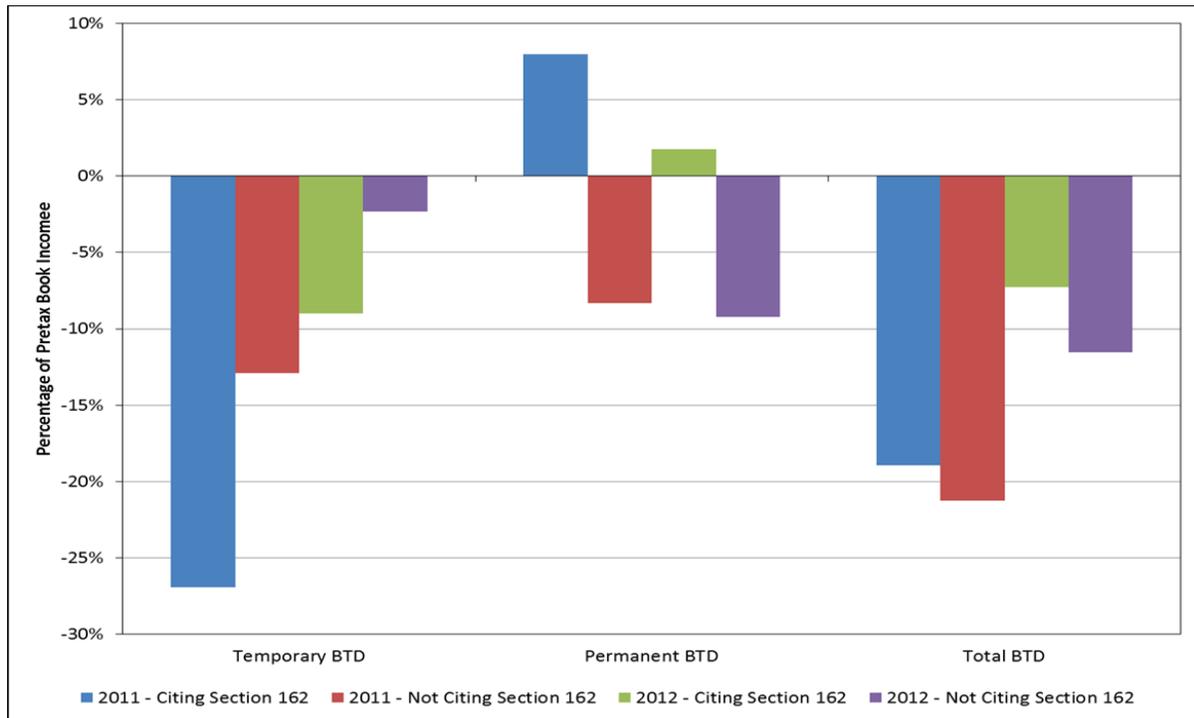


FIGURE 6C. 2011–2012 U.S. Corporation M-3: 2-Year Average Total BTD as Percentage of Adjusted Total Book Income for Top 6 Lines—UTP Filers Citing/Not Citing IRC Section 162 (SEC/10K Public Financial Statements)

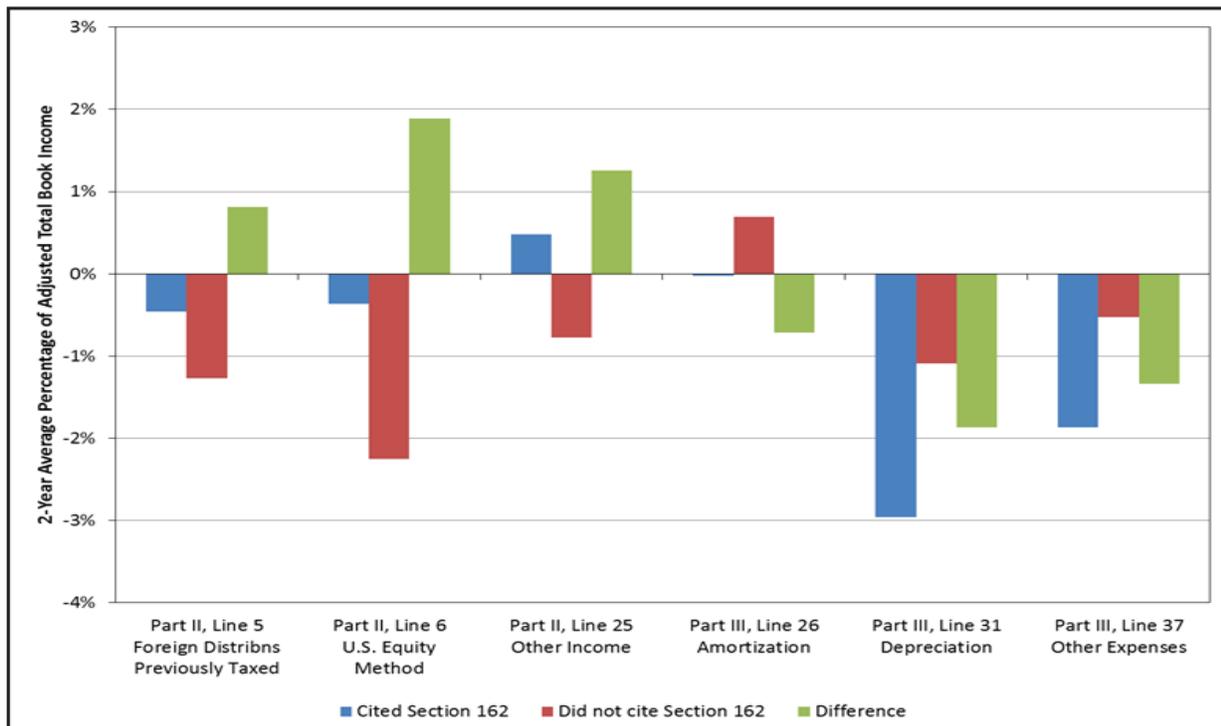


Figure 6C reports total BTD for six specific Schedule M-3 lines for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 162.⁵⁴ The total BTD for each of the six lines are expressed as a percentage of adjusted total income and averaged for 2011 and 2012. The lines are the top (that is, the most extreme) six lines on Schedule M-3, Parts II and III, in terms of the absolute value of the difference in 2-year-average scaled BTD between citers and noncitters.⁵⁵ The percentages for each of these lines are provided below.

Schedule M-3 Lines Featured in Figure 6C		Citers	Noncitters	Difference
Part II, line 5:	Exclusion in tax income of previously taxed foreign distributions	-0.46%	-1.27%	0.81%
Part II, line 6:	Exclusion from tax income of U.S. equity method income	-0.36%	-2.25%	1.89%
Part II, line 25:	Adjustment to other income with difference in tax income	0.48%	-0.77%	1.25%
Part III, line 26:	Adjustment to amortization/impairment of goodwill expense/deduction in tax income	-0.02%	0.69%	-0.72%
Part III, line 31:	Adjustment to depreciation expense/deduction in tax income	-2.96%	-1.09%	-1.86%
Part III, line 37:	Adjustment to other expense/deduction with difference in tax income	-1.87%	-0.53%	-1.34%

5. IRC Section 199 Citers

Schedule UTP filers are required to list one to three applicable IRC sections for each listed UTP concise description. Schedule UTP, Part I, tabulates the IRC information for each current year UTP concise description listed in Schedule UTP, Part III. SOI tabulates the first 10 rows of data on Schedule UTP, Part I. The five IRC sections appearing most frequently in the SOI file for Tax Years 2011 and 2012 are: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized costs). In 2012, seventy-two percent of such filers cited at least one of the five IRC sections. We categorize Schedule UTP filers as citing or not citing a specific IRC section. In this part of our study we focus on data describing Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 199.⁵⁶

Tables 7A and 7B and Figures 7A and 7B present 2012 data describing characteristics of Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 199. Figures 7A and 7B include 2011 data for comparison. Figure 7C presents 2011 and 2012 average data.

Schedule UTP filers citing IRC section 199 are, with the exception of total assets, generally larger than noncitters when study is restricted to the 1,176 Schedule UTP filers in 2012 with SEC 10K/Public financial statements and \$100 million or more in assets. The mean asset size as reported on the Form 1120, Schedule L, by such Schedule UTP filers is \$17,347 million for the 130 filers citing IRC section 199 and \$21,679 million for the 1,046 filers not citing IRC section 199 (see Table 7B). The mean worldwide financial net income of the citers (reported on Schedule M-3, Part I, line 4) is \$1,483 million compared to \$437 million for the noncitters. The mean foreign nonincludible income of citers (reported on Schedule M-3, Part I, line 5) is -\$1,267 million (shown as negative since the income is removed in calculating book income) compared to -\$383 million for nonfilers. After the required Schedule M-3, Part I, adjustments, the mean book income for citers (reported on Schedule M-3, Part I, line 11) is \$764 million compared to \$435 million for the noncitters. Adding back U.S. Federal tax expense, mean pretax book income is \$1,140 million for the citers compared to \$513 million for the

⁵⁴ The six lines are selected from and represent approximately 10 percent of the 68 specified and other-with-difference lines on Schedule M-3, Parts II and III. See the discussion and footnotes for Figure 3C.

⁵⁵ For the current study, we exclude interest income (Part II, line 13) and interest expense (Part III, line 8) as key lines because of a data anomaly in the 2011 and 2012 data. See the discussion and footnotes for Figure 3C.

⁵⁶ We focus on Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more but with Audited (but not SEC 10K/Public) or Unaudited financial statements.

TABLE 7A. 2012 U.S. Corporation 1120 Schedule M-3: Mini M-3—IRC 199 Cited: SEC 10K/Public: Assets \$100 Million or More

	\$ Million				Percentage of Adjusted Total Income				Percentage of Pretax Book					
	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Column A Book	Column B Temporary	Column C Permanent	Column D Tax	Column B Temporary	Column C Permanent	Column D Tax	Column B Temporary	Column C Permanent	Column D Tax
	Total	Difference	Total	Difference	Total	Difference	Total	Difference	Total	Difference	Total	Difference	Total	Difference
UTP Filers Citing IRC 199														
Adj. Other income—no difference	1,429,379	0	0	1,429,379	0	206.12%	0.00%	206.12%	0.00%	0.00%	206.12%	0.00%	0.00%	0.00%
Adjusted COGS	-903,090	-1,421	-390	-904,902	-1,812	-130.23%	-0.20%	-130.49%	-0.26%	-0.96%	-130.49%	-0.26%	-0.26%	-1.22%
Adjusted gross profit	526,289	-1,421	-390	524,477	-1,812	75.89%	-0.20%	75.63%	-0.26%	-0.96%	75.63%	-0.26%	-0.26%	-1.22%
Specified income	94,921	-9,698	14,867	100,090	5,169	13.69%	-1.40%	14.43%	0.75%	-6.54%	14.43%	10.03%	10.03%	3.49%
Adj. Other income—difference	72,261	-1,339	-2,722	68,199	-4,062	10.42%	-0.19%	9.83%	-0.59%	-0.90%	9.83%	-1.84%	-1.84%	-2.74%
Adjusted total income	693,471	-12,458	11,755	692,766	-705	100.00%	-1.80%	99.90%	-0.10%	-8.41%	99.90%	7.93%	7.93%	-0.48%
Specified expense/deduction	-162,336	4,489	-8,080	-165,927	-3,590	-23.41%	0.65%	-23.93%	-0.52%	3.03%	-23.93%	-5.45%	-5.45%	-2.42%
Other exp./ded.—difference	-225,980	-8,006	-2,375	-236,362	-10,381	-32.59%	-1.15%	-34.08%	-1.50%	-5.40%	-34.08%	-1.60%	-1.60%	-7.00%
Other exp./ded.—no difference	-156,951	0	0	-156,951	0	-22.63%	0.00%	-22.63%	0.00%	0.00%	-22.63%	0.00%	0.00%	0.00%
Pretax net income	148,204	-15,975	1,300	133,526	-14,676	21.37%	-2.30%	19.25%	-2.12%	-10.78%	19.25%	0.88%	0.88%	-9.90%
UTP Filers Not Citing IRC 199														
Adj. Other income—no difference	5,186,232	0	0	5,186,232	0	172.93%	0.00%	172.93%	0.00%	0.00%	172.93%	0.00%	0.00%	0.00%
Adjusted COGS	-3,167,989	-5,585	189	-3,173,384	-5,396	-105.64%	-0.19%	-105.82%	-0.18%	-1.04%	-105.82%	-0.04%	-0.04%	-1.00%
Adjusted gross profit	2,018,243	-5,585	189	2,012,848	-5,396	67.30%	-0.19%	67.12%	-0.18%	-1.04%	67.12%	0.04%	0.04%	-1.00%
Specified income	904,390	39,604	-47,129	896,735	-7,525	30.16%	1.32%	29.90%	-0.25%	7.37%	29.90%	-8.78%	-8.78%	-1.40%
Adj. Other income—difference	76,357	25,449	-14,010	87,796	11,439	2.55%	0.85%	2.93%	0.38%	4.74%	2.93%	-2.61%	-2.61%	2.13%
Adjusted total income	2,998,990	59,468	-60,950	2,997,379	-1,482	100.00%	1.98%	99.95%	-0.05%	11.07%	99.95%	-11.35%	-11.35%	-0.28%
Specified expense/deduction	-824,668	-40,249	16,325	-848,463	-23,925	-27.50%	-1.34%	-28.29%	-0.80%	-7.49%	-28.29%	3.04%	3.04%	-4.46%
Other exp./ded.—difference	-864,570	-32,770	2,494	-894,846	-30,276	-28.83%	-1.09%	-29.84%	-1.01%	-6.10%	-29.84%	0.46%	0.46%	-5.64%
Other exp./ded.—no difference	-772,722	0	0	-772,722	0	-25.77%	0.00%	-25.77%	0.00%	0.00%	-25.77%	0.00%	0.00%	0.00%
Pretax net income	537,030	-13,551	-42,131	481,348	-55,683	17.91%	-0.45%	16.05%	-1.86%	-2.52%	16.05%	-7.85%	-7.85%	-10.37%

nonciters. Citers introduce mean BTD of -\$113 million to adjust pretax book to mean tax net income (reported on Schedule M-3, Part II, line 30, column D, and on Form 1120, page 1, line 28) of \$1,027 million compared to nonciters that introduce mean BTD of -\$53 million to adjust pretax book to mean tax net income of \$460 million.⁵⁷

We scaled our initial aggregate dollar data by the adjusted-total-income measure we developed for our Mini M-3 analysis to make data for filers that cite or do not cite a specific IRC section more comparable. We used the adjusted-total-income book amount as a common-size scaling factor and compared percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers that cite or do not cite a specific IRC section, total pretax income BTD is expressed as a percentage of total pretax book income (see Tables 7A and 7B).

Table 7B shows that Schedule UTP filers in 2012 having SEC 10K/Public financial statements and \$100 million or more in assets, and citing IRC section 199 report:

- more worldwide income scaled as a percentage of adjusted total income than similar filers not citing IRC section 199 (27.80 percent compared to 15.24 percent);
- more scaled foreign nonincludible income (-23.75 percent versus -13.37 percent);
- *less* scaled book income (14.31 percent versus 15.19 percent);
- more scaled pretax book income (21.37 percent versus 17.91 percent);
- more scaled tax net income (19.25 percent versus 16.05 percent); and
- more negative scaled BTD (-2.12 percent versus -1.86 percent).

If the BTD is scaled by pretax book income, the filers citing IRC section 199 reduce pretax book income by -9.90 percent to determine tax net income compared to a reduction of -10.37 percent by nonciters (see Table 7A, last row, last column, in each panel).

Figure 7A compares Schedule M-3 data as percentages of adjusted total income for 2011 and 2012 Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 199, for seven key items: worldwide financial income (Part I, line 4); foreign nonincludible income (Part I, line 5); adjustments to eliminations (Part I, line 8); other adjustments (Part I, line 10); book income (Part I, line 11); pretax book income; and tax net income. With the exception of adjustments to eliminations for 2011 and 2012 and book income for 2012, for both years, filers citing IRC section 199 report larger scaled amounts for the seven items than nonciters.

Figure 7B uses pretax book income for scaling 2011 and 2012 BTD, temporary, permanent, and total, for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 199. In 2011, citers have less negative temporary BTD than nonciters, in 2012 citers have more negative temporary BTD than nonciters, and in both 2011 and 2012 citers have positive permanent BTD compared to negative permanent BTD for nonciters. In aggregate, citers of IRC section 199 have less negative total BTD than nonciters in both 2011 and 2012. Stated differently, the otherwise similar nonciters use negative total BTD to reduce tax net income more than the citers.

⁵⁷ Negative total BTD adjustments added to pretax book income result in tax net income that is lower than pretax book income.

**TABLE 7B. 2012 U.S. Corporation 1120 Schedule M-3: M-3, Part I—IRC 199 Cited: SEC 10K/
Public: Assets \$100 Million or More**

	Mean Amount Reported (\$ Million)		
	IRC 199 Cited	IRC 199 Not Cited	Difference
Worldwide financial net income	1,483	437	1,046
(Foreign nonincludible income)	-1,267	-383	-884
(U.S. nonincludible income)	-56	-29	-28
Other includible income	-3	1	-5
Adjustments to eliminations	336	289	47
Other adjustments	273	113	160
Book net income includible corp.	764	435	328
Pretax net income—book	1,140	513	627
Tax net income	1,027	460	567
Total BTD difference	-113	-53	-60
Total assets—Schedule L	17,347	21,679	-4,332
Assets—Financial statements	16,991	19,228	-2,237
Assets—Foreign nonincludible	8,147	7,500	648
Assets—U.S. nonincludible	727	643	83
Assets—Other includible	32	128	-97
	Percentage of Adjusted Total Income Book		
	IRC 199 Cited	IRC 199 Not Cited	Difference
Worldwide financial net income	27.80%	15.24%	12.55%
(Foreign nonincludible income)	-23.75%	-13.37%	-10.38%
(U.S. nonincludible income)	-1.05%	-1.00%	-0.06%
Other includible income	-0.06%	0.05%	-0.11%
Adjustments to eliminations	6.30%	10.09%	-3.79%
Other adjustments	5.12%	3.94%	1.19%
Book net income includible corp.	14.31%	15.19%	-0.87%
Pretax net income—book	21.37%	17.91%	3.46%
Tax net income	19.25%	16.05%	3.20%
Total BTD difference	-2.12%	-1.86%	-0.26%
	Distribution of Returns		
	IRC 199 Cited	IRC 199 Not Cited	Total
Number of returns	130	1,046	1,176
Share of returns	11.1%	88.9%	100.0%

In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

Figure 7C reports total BTD for six specific Schedule M-3 lines for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 199.⁵⁸ The total BTD for each of the six lines are expressed as a percentage of adjusted total income and averaged for 2011 and

⁵⁸ The six lines are selected from and represent approximately 10 percent of the 68 specified and other-with-difference lines on Schedule M-3, Parts II and III. See the discussion and footnotes for Figure 3C.

FIGURE 7A. 2011–2012 U.S. Corporation Key M-3 Data as Percentage of Adjusted Total Book Income for SEC/10K Public Financial Statements by IRC Section 199

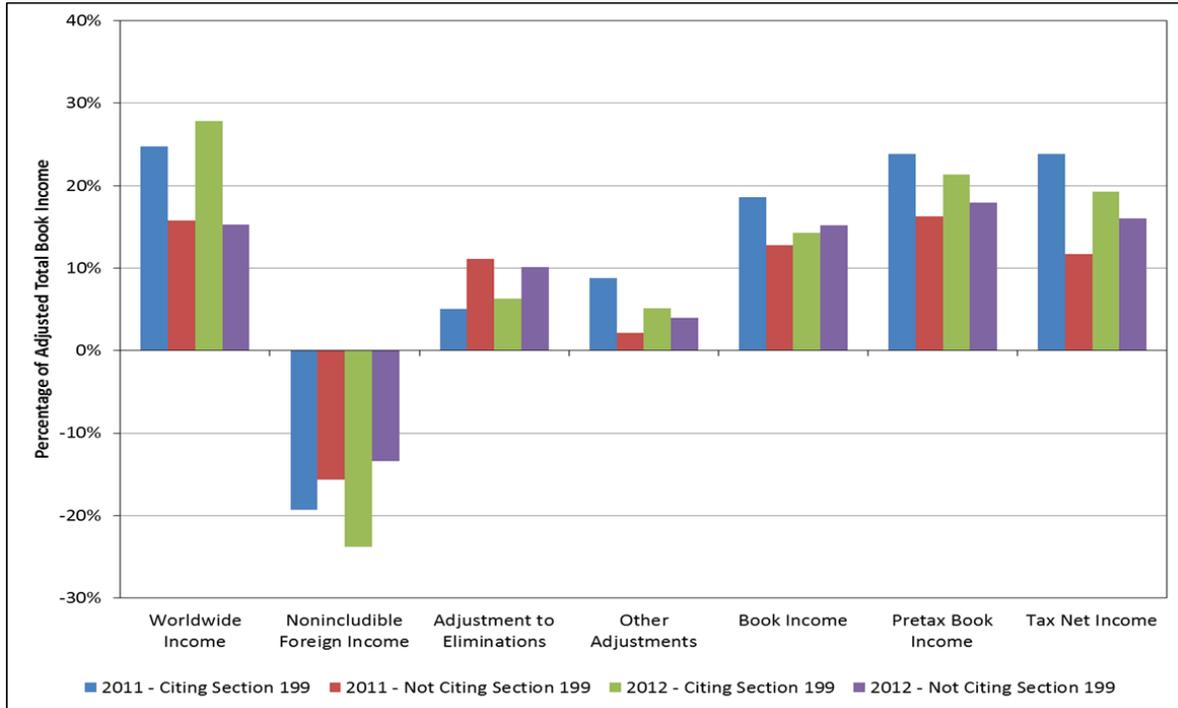


FIGURE 7B. 2011–2012 U.S. Corporation M-3: BTD as Percentage of Pretax Book Income for SEC/10K Public Financial Statements by IRC Section 199

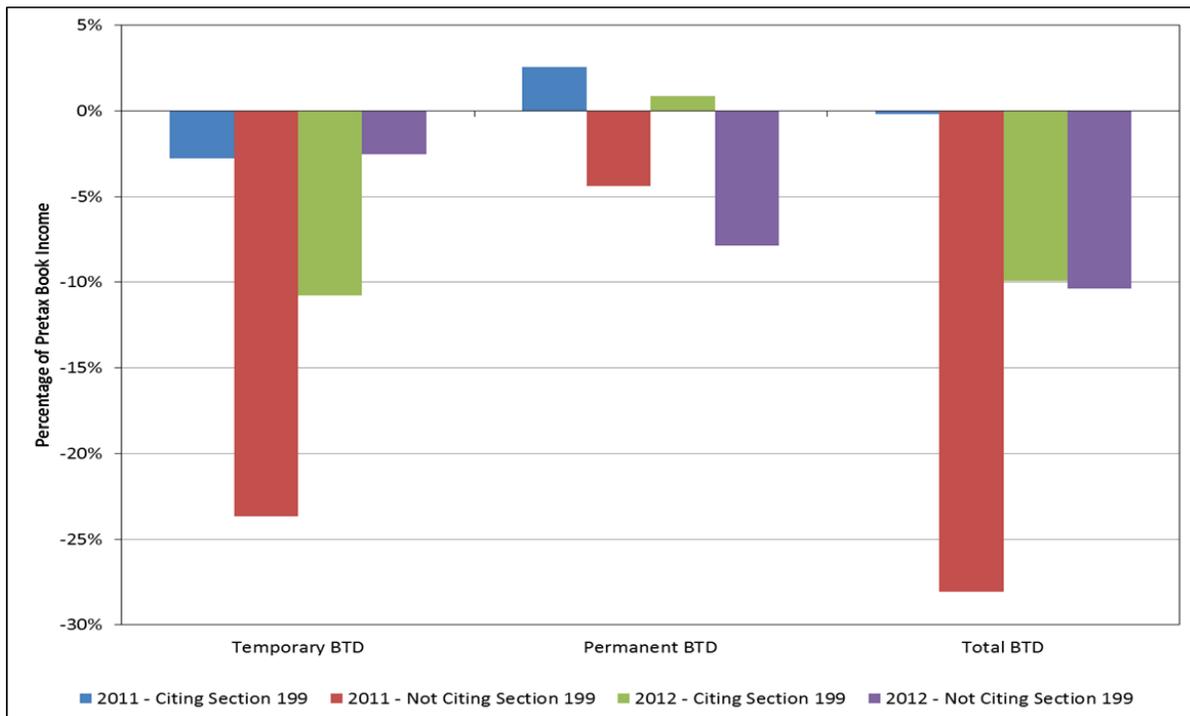
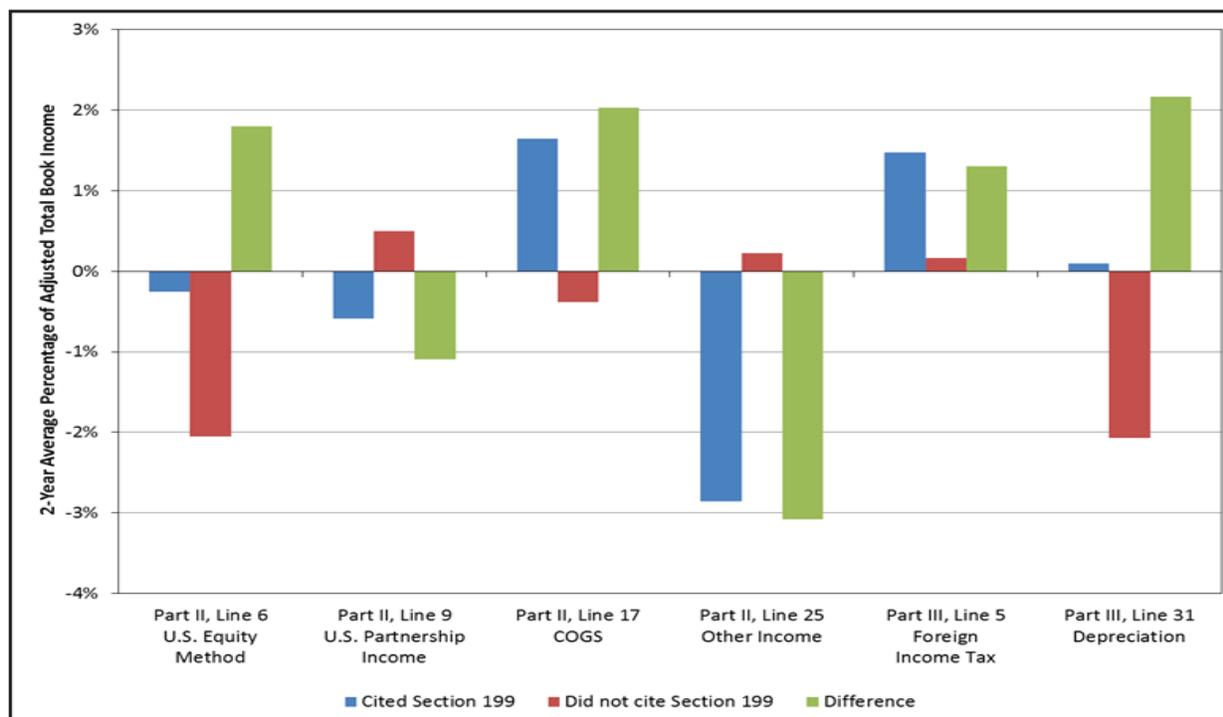


FIGURE 7C. 2011–2012 U.S. Corporation M-3: 2-Year Average Total BTD as Percentage of Adjusted Total Book Income for Top 6 Lines—UTP Filers Citing/Not Citing IRC Section 199 (SEC/10K Public Financial Statements)



2012. The lines are the top (that is, the most extreme) six lines on Schedule M-3, Parts II and III, in terms of the absolute value of the difference in 2-year-average scaled BTD between citers and noncitters.⁵⁹ The percentages for each of these lines are provided below.

Schedule M-3 Lines Featured in Figure 7C		Citers	Noncitters	Difference
Part II, line 6:	Exclusion from tax income of U.S. equity method income	-0.25%	-2.05%	1.80%
Part II, line 9:	Adjustments to U.S. partnership income to include all Schedule K-1 income in tax income	-0.59%	0.50%	-1.09%
Part II, line 17:	Adjustments to COGS in tax income	1.65%	-0.38%	2.03%
Part II, line 25:	Adjustment to other income with difference in tax income	-2.86%	0.22%	-3.08%
Part III, line 5:	Adjustment to foreign income tax expense/deduction in tax income	1.47%	0.17%	1.31%
Part III, line 31:	Adjustment to depreciation expense/deduction in tax income	0.09%	-2.07%	2.17%

6. IRC Section 263 Citers

Schedule UTP filers are required to list one to three applicable IRC sections for each listed UTP concise description. Schedule UTP, Part I, tabulates the IRC information for each current year UTP concise description listed in Schedule UTP, Part III. SOI tabulates the first 10 rows of data on Schedule UTP, Part I. The five IRC sections appearing most frequently in the SOI file for Tax Years 2011 and 2012 are: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized costs). In 2012, seventy-two percent of such filers cited at least one of the five IRC sections. We categorize Schedule UTP filers as citing or not citing a specific IRC section. In this part of our study we focus

⁵⁹ For the current study, we exclude interest income (Part II, line 13) and interest expense (Part III, line 8) as key lines because of a data anomaly in the 2011 and 2012 data. See the discussion and footnotes for Figure 3C.

on data describing Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 263.⁶⁰

Tables 8A and 8B and Figures 8A and 8B present 2012 data describing characteristics of Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 263. Figures 8A and 8B include 2011 data for comparison. Figure 8C presents 2011 and 2012 average data.

Schedule UTP filers citing IRC section 263 are generally larger than nonciters when study is restricted to the 1,176 Schedule UTP filers in 2012 with SEC 10K/Public financial statements and \$100 million or more in assets. The mean asset size as reported on the Form 1120, Schedule L, by such Schedule UTP filers is \$25,325 million for the 91 filers citing IRC section 263 and \$20,854 million for the 1,085 filers not citing IRC section 263 (see Table 8B). The mean worldwide financial net income of the citers (reported on Schedule M-3, Part I, line 4) is \$1,004 million compared to \$515 million for the nonciters. The mean foreign nonincludible income of citers (reported on Schedule M-3, Part I, line 5) is -\$703 million (shown as negative since the income is removed in calculating book income) compared to -\$463 million for nonfilers. After the required Schedule M-3, Part I, adjustments, the mean book income for citers (reported on Schedule M-3, Part I, line 11) is \$1,384 million compared to \$395 million for the nonciters. Adding back U.S. Federal tax expense, mean pretax book income is \$1,615 million for the citers compared to \$496 million for the nonciters. Citers introduce mean BTD of -\$258 million to adjust pretax book to mean tax net income (reported on Schedule M-3, Part II, line 30, column D, and on Form 1120, page 1, line 28) of \$1,357 million compared to nonciters that introduce mean BTD of -\$43 million to adjust pretax book to mean tax net income of \$453 million.⁶¹

We scale our initial aggregate dollar data by the adjusted-total-income measure we develop for our Mini M-3 analysis to make data for filers that cite or do not cite a specific IRC section more comparable. We used the adjusted-total-income book amount as a common-size scaling factor and compare percentages of adjusted-total-income book to remove or minimize the impact of differences in the size of corporations from our analysis. In addition, for the purposes of comparing the Schedule M-3 characteristics of Schedule UTP filers that cite or do not cite a specific IRC section, total pretax income BTD is expressed as a percentage of total pretax book income (see Tables 8A and 8B).

Table 8B shows that Schedule UTP filers in 2012 having SEC 10K/Public financial statements and \$100 million or more in assets, and citing IRC section 263 report:

- *less* worldwide income scaled as a percentage of adjusted total income than similar filers not citing IRC section 263 (15.44 percent compared to 18.01 percent);
- *less* scaled foreign nonincludible income (-10.80 percent versus -16.19 percent);
- *more* scaled book income (21.27 percent versus 13.83 percent);
- *more* scaled pretax book income (24.82 percent versus 17.36 percent);
- *more* scaled tax net income (20.85 percent versus 15.85 percent); and
- *more* negative scaled BTD (-3.97 percent versus -1.51 percent).

If the BTD is scaled by pretax book income, the filers citing IRC section 263 reduce pretax book income by -15.99 percent to determine tax net income compared to a reduction of -8.71 percent by nonciters (see Table 8A, last row, last column, in each panel).

⁶⁰ We focus on Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets because of the size and economic importance of such corporations and because of the comparatively low number of Schedule UTP filers with assets of less than \$100 million or with assets of \$100 million or more but with Audited (but not SEC 10K/Public) or Unaudited financial statements.

⁶¹ Negative total BTD adjustments added to pretax book income result in tax net income that is lower than pretax book income.

TABLE 8A. 2012 U.S. Corporation 1120 Schedule M-3: Mini M-3—IRC 263 Cited: SEC 10K/Public: Assets \$100 Million or More

	\$ Million					Percentage of Adjusted Total Income					Percentage of Pretax Book				
	Column A	Column B	Column C	Column D	Total	Column A	Column B	Column C	Column D	Total	Column A	Column B	Column C	Total	
	Book	Temporary	Permanent	Tax	Difference	Book	Temporary	Permanent	Tax	Difference	Temporary	Permanent	Permanent	Difference	
UTP Filers Citing IRC 263															
Adj. Other income—no difference	1,249,202	0	0	1,249,202	0	210.97%	0.00%	0.00%	210.97%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Adjusted COGS	-792,935	-3,996	-177	-797,108	-4,173	-133.91%	-0.67%	-0.03%	-134.62%	-0.70%	-2.72%	-0.12%	-0.12%	-2.84%	-2.84%
Adjusted gross profit	456,267	-3,996	-177	452,094	-4,173	77.05%	-0.67%	-0.03%	76.35%	-0.70%	-2.72%	-0.12%	-0.12%	-2.84%	-2.84%
Specified income	110,524	-13	-1,665	108,846	-1,678	18.67%	0.00%	-0.28%	18.38%	-0.28%	-0.01%	-1.13%	-1.13%	-1.14%	-1.14%
Adj. Other income—difference	25,346	4,309	-5,274	24,381	-965	4.28%	0.73%	-0.89%	4.12%	-0.16%	2.93%	-3.59%	-3.59%	-0.66%	-0.66%
Adjusted total income	592,137	300	-7,116	585,321	-6,816	100.00%	0.05%	-1.20%	98.85%	-1.15%	0.20%	-4.84%	-4.84%	-4.64%	-4.64%
Specified expense/deduction	-146,298	-5,693	3,977	-148,014	-1,716	-24.71%	-0.96%	0.67%	-25.00%	-0.29%	-3.87%	2.71%	2.71%	-1.17%	-1.17%
Other exp./ded.—difference	-153,135	-17,098	2,130	-168,103	-14,968	-25.86%	-2.89%	0.36%	-28.39%	-2.53%	-11.63%	1.45%	1.45%	-10.18%	-10.18%
Other exp./ded.—no difference	-145,715	0	0	-145,715	0	-24.61%	0.00%	0.00%	-24.61%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Pretax net income	146,989	-22,491	-1,009	123,489	-23,500	24.82%	-3.80%	-0.17%	20.85%	-3.97%	-15.30%	-0.69%	-0.69%	-15.99%	-15.99%
UTP Filers Not Citing IRC 263															
Adj. Other income—no difference	5,366,410	0	0	5,366,410	0	173.09%	0.00%	0.00%	173.09%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Adjusted COGS	-3,278,144	-3,010	-25	-3,281,178	-3,035	-105.74%	-0.10%	0.00%	-105.83%	-0.10%	-0.56%	0.00%	0.00%	-0.56%	-0.56%
Adjusted gross profit	2,088,266	-3,010	-25	2,085,232	-3,035	67.36%	-0.10%	0.00%	67.26%	-0.10%	-0.56%	0.00%	0.00%	-0.56%	-0.56%
Specified income	888,787	29,919	-30,597	887,979	-678	28.67%	0.97%	-0.99%	28.64%	-0.02%	5.56%	-5.68%	-5.68%	-0.13%	-0.13%
Adj. Other income—difference	123,270	19,801	-11,458	131,613	8,343	3.98%	0.64%	-0.37%	4.25%	0.27%	3.68%	-2.13%	-2.13%	1.55%	1.55%
Adjusted total income	3,100,323	46,710	-42,080	3,104,824	4,630	100.00%	1.51%	-1.36%	100.15%	0.15%	8.68%	-7.82%	-7.82%	0.86%	0.86%
Specified expense/deduction	-840,707	-30,067	4,268	-866,376	-25,799	-27.12%	-0.97%	0.14%	-27.94%	-0.83%	-5.59%	0.79%	0.79%	-4.79%	-4.79%
Other exp./ded.—difference	-937,415	-23,678	-2,011	-963,104	-25,689	-30.24%	-0.76%	-0.06%	-31.06%	-0.83%	-4.40%	-0.37%	-0.37%	-4.77%	-4.77%
Other exp./ded.—no difference	-783,958	0	0	-783,958	0	-25.29%	0.00%	0.00%	-25.29%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Pretax net income	538,243	-7,035	-39,823	491,386	-46,858	17.36%	-0.23%	-1.28%	15.85%	-1.51%	-1.31%	-7.40%	-7.40%	-8.71%	-8.71%

TABLE 8B. 2012 U.S. Corporation 1120 Schedule M-3: M-3, Part I—IRC 263 Cited: SEC 10K/Public: Assets \$100 Million or More

	Mean Amount Reported (\$ Million)		
	IRC 263 Cited	IRC 263 Not Cited	Difference
Worldwide financial net income	1,004	515	490
(Foreign nonincludible income)	-703	-463	-240
(U.S. nonincludible income)	-35	-31	-4
Other includible income	3	1	3
Adjustments to eliminations	-25	321	-346
Other adjustments	1,137	46	1,091
Book net income includible corp.	1,384	395	989
Pretax net income—book	1,615	496	1,119
Tax net income	1,357	453	904
Total BTD difference	-258	-43	-215
Total assets—Schedule L	25,325	20,854	4,471
Assets—Financial statements	25,864	18,403	7,461
Assets—Foreign nonincludible	3,816	7,886	-4,070
Assets—U.S. nonincludible	722	647	75
Assets—Other includible	345	98	247
	Percentage of Adjusted Total Income Book		
	IRC 263 Cited	IRC 263 Not Cited	Difference
Worldwide financial net income	15.44%	18.01%	-2.58%
(Foreign nonincludible income)	-10.80%	-16.19%	5.39%
(U.S. nonincludible income)	-0.54%	-1.10%	0.56%
Other includible income	0.05%	0.03%	0.03%
Adjustments to eliminations	-0.38%	11.24%	-11.62%
Other adjustments	17.47%	1.62%	15.86%
Book net income includible corp.	21.27%	13.83%	7.44%
Pretax net income—book	24.82%	17.36%	7.46%
Tax net income	20.85%	15.85%	5.01%
Total BTD difference	-3.97%	-1.51%	-2.46%
	Distribution of Returns		
	IRC 263 Cited	IRC 263 Not Cited	Total
Number of returns	91	1,085	1,176
Share of returns	7.7%	92.3%	100.0%

Figure 8A compares Schedule M-3 data as percentages of adjusted total income for 2011 and 2012 Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 263, for seven key items: worldwide financial income (Part I, line 4); foreign nonincludible income (Part I, line 5); adjustments to eliminations (Part I, line 8); other adjustments (Part I, line 10); book income (Part I, line 11); pretax book income; and tax net income. With the exception of worldwide income for 2012, foreign nonincludible income for 2011 and 2012, and adjustments to eliminations for 2011 and 2012, for both years, filers citing IRC section 263 report larger scaled amounts for the seven items than nonciters.

Figure 8B uses pretax book income for scaling 2011 and 2012 BTD, temporary, permanent, and total, for Schedule UTP filers, with SEC 10K/Public financial statements and \$100 million or more in assets, citing or not citing IRC section 263. In 2011 and 2012, citers have more negative temporary BTD than nonciters. In 2011 citers have positive permanent BTD and in 2012 less negative permanent BTD compared to negative permanent BTD for nonciters in both years. In aggregate, citers of IRC section 263 have more negative total BTD than nonciters in both 2011 and 2012. Stated differently, the otherwise similar nonciters use negative total BTD to reduce tax net income less than the citers.

FIGURE 8A. 2011–2012 U.S. Corporation Key M-3 Data as Percentage of Adjusted Total Book Income for SEC/10K Public Financial Statements by IRC Section 263

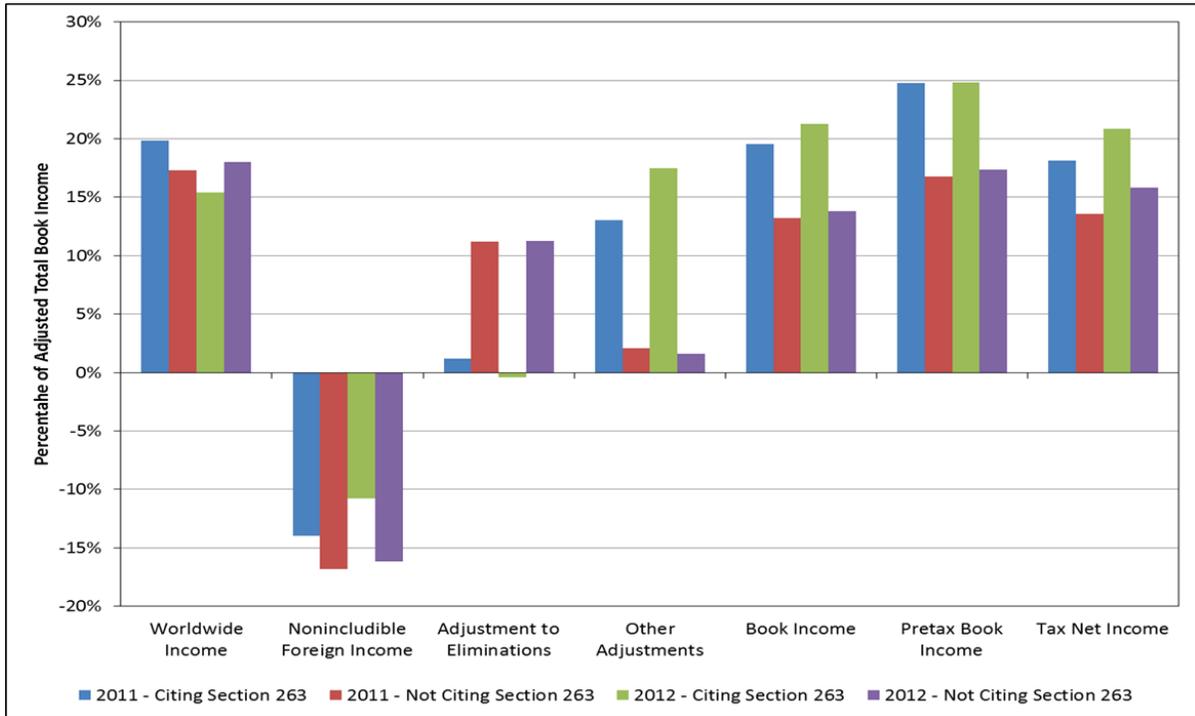


FIGURE 8B. 2011–2012 U.S. Corporation M-3: BTD as Percentage of Pretax Book Income for SEC/10K Public Financial Statements by IRC Section 263

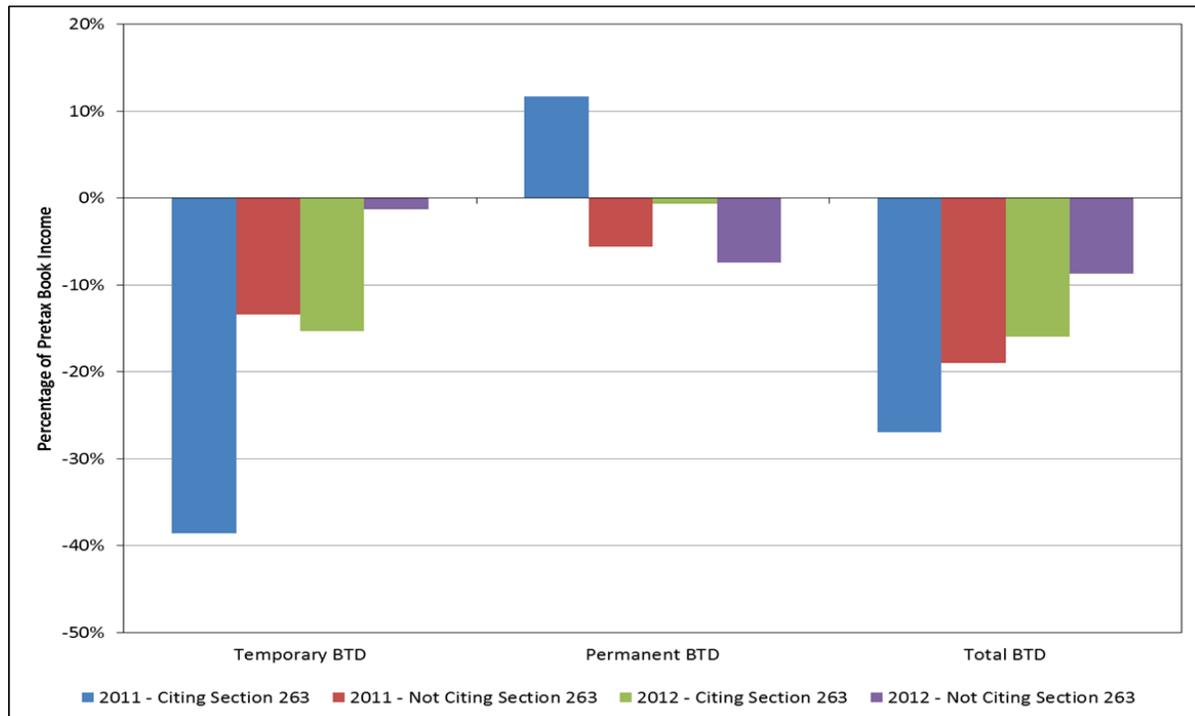
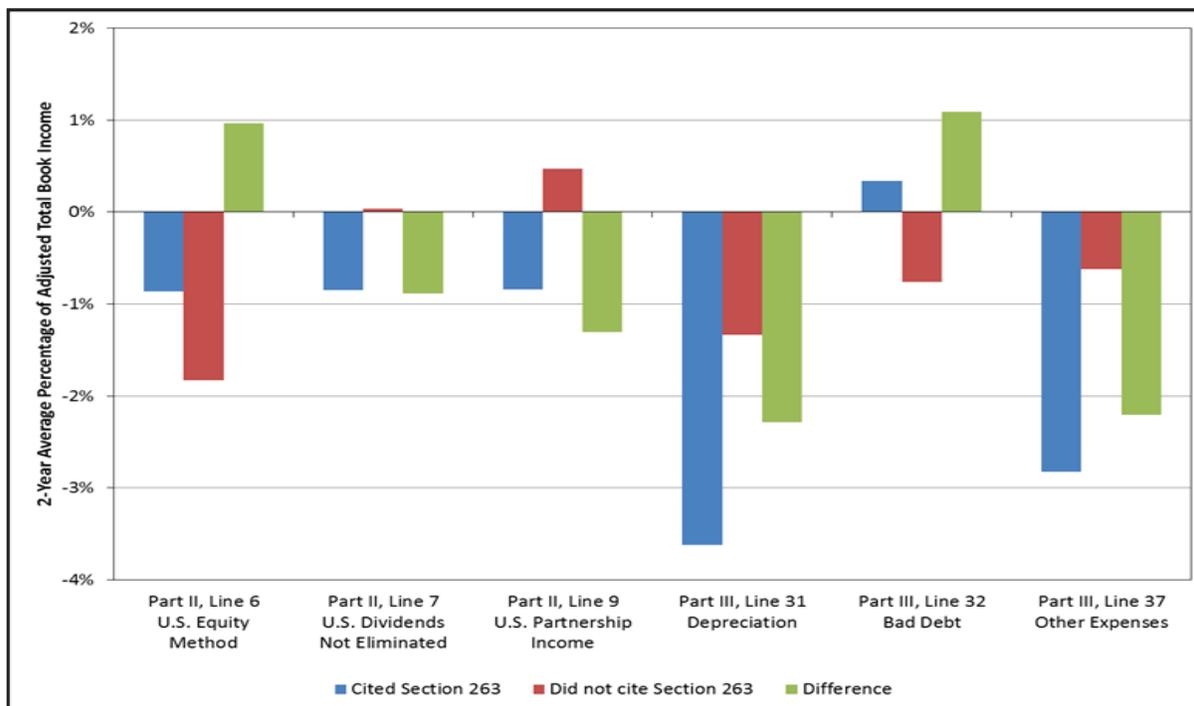


FIGURE 8C. 2011–2012 U.S. Corporation M-3: 2-Year Average Total BTD as Percentage of Adjusted Total Book Income for Top 6 Lines—UTP Filers Citing/Not Citing IRC Section 263 (SEC/10K Public Financial Statements)



In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

Figure 8C reports total BTD for six specific Schedule M-3 lines for Schedule UTP filers, with SEC 10K/ Public financial statements and \$100 million or more in assets, citing or not citing IRC section 263.⁶² The total BTD for each of the six lines are expressed as a percentage of adjusted total income and averaged for 2011 and 2012. The lines are the top (that is, the most extreme) six lines on Schedule M-3, Parts II and III, in terms of the absolute value of the difference in 2-year-average scaled BTD between citers and noncitters.⁶³ The percentages for each of these lines are provided below.

Schedule M-3 Lines Featured in Figure 8C		Citers	Noncitters	Difference
Part II, line 6:	Exclusion from tax income of U.S. equity method income	-0.86%	-1.83%	0.97%
Part II, line 7:	Adjustments to U.S. dividends, not eliminated in consolidation, in tax income	-0.85%	0.04%	-0.89%
Part II, line 9:	Adjustments to U.S. partnership income to include all Schedule K-1 income in tax income	-0.84%	0.47%	-1.31%
Part III, line 31:	Adjustment to depreciation expense/deduction in tax income	-3.62%	-1.33%	-2.29%
Part III, line 32:	Adjustments to bad debt expense/deduction recognition in tax income	0.34%	-0.76%	1.09%
Part III, line 37:	Adjustment to other expense/deduction with difference in tax income	-2.82%	-0.62%	-2.21%

⁶² The six lines are selected from and represent approximately 10 percent of the 68 specified and other-with-difference lines on Schedule M-3, Parts II and III. See the discussion and footnotes for Figure 3C.

⁶³ For the current study, we exclude interest income (Part II, line 13) and interest expense (Part III, line 8) as key lines because of a data anomaly in the 2011 and 2012 data. See the discussion and footnotes for Figure 3C.

Part IV. Summary and Conclusions

The study compares 2011–2012 Schedule M-3 and Form 1120 tax return data profiles for Schedule UTP filers and nonfilers. For Schedule UTP filers with SEC 10K/Public financial statements and \$100 million or more in assets, the study further compares Schedule M-3 profiles of corporations that cite or do not cite on Schedule UTP any of the five most commonly cited IRC sections: 482 (transfer pricing); 41 (research credit); 162 (trade or business expense); 199 (domestic production activities deduction); and 263 (capitalized costs). In 2012, seventy-two percent of such filers cited at least one of the five IRC sections.

In developing filters and quantitative models for return selection it is useful to identify data items that are effective in separating or distinguishing between otherwise similar groups of returns that, in fact, have different underlying characteristics of interest. These separating or distinguishing data items may, for example, be data items with extreme absolute differences in the average values of the data items for the groups to be separated. The final quantitative models may take many different forms in using the identified data items with extreme absolute differences. The models would generally operate at the microdata level and generally use both the sign and amount of the data items.

The authors conclude that Schedule UTP filers and nonfilers and Schedule UTP filers that cite frequently cited IRC sections have Schedule M-3 data profiles that are sufficiently different that Schedule M-3 quantitative models could be developed to detect on Schedule UTP nonfiler returns the underlying issues related to the IRC sections cited by Schedule UTP filers. The authors believe such models may assist LB&I return selection.

References

- Boynton, Charles, Portia DeFilippes, and Ellen Legel. "Prelude to Schedule M-3: Schedule M-1 Corporate Book-Tax Difference Data 1990-2003." *Tax Notes* 109, No. 12 (December 19, 2005): 1579–99.
- Boynton, Charles, Portia DeFilippes, and Ellen Legel. "Distribution of Schedule M-1 Corporate Book-Tax Difference Data 1990-2003 for Three Large-Size and Three Industry Groups." *Tax Notes* 111, No. 2 (April 10, 2006a): 177–212.
- Boynton, Charles, Portia DeFilippes, and Ellen Legel. "A First Look at 2004 Schedule M-3 Reporting by Large Corporations." *Tax Notes* 112, No. 11 (September 11, 2006b): 943–981. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Boynton, Charles, Portia DeFilippes, and Ellen Legel. "A First Look at 2005 Schedule M-3 Corporate Reporting." *Tax Notes*, 121, No.5 (November 3, 2008): p. 563. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Boynton, Charles, Portia DeFilippes, Ellen Legel, and Todd Reum. "A First Look at 2007 Schedule M-3 Reporting by Large Corporations." *Tax Notes* 132, No. 7 (August 15, 2011): p. 689. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Boynton, Charles, Portia DeFilippes, Ellen Legel, and Todd Reum. "A First Look at 2010 Schedule M-3 Reporting and Schedule UTP." *Tax Notes*, 144, No.3 (July 21, 2014): p. 253. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Boynton, Charles, Portia DeFilippes, Ellen Legel, and Lisa Rupert. "2010-2011 Schedule M-3 Profiles and Schedule UTP Filing Status." *Tax Notes*, 145, No.5 (November 3, 2014): p. 535. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Boynton, Charles, and Barbara Livingston. "Partnerships with Reportable Entity Partners." *Tax Notes* 128, No. 9 (August 30, 2010): 949–958. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Boynton, Charles, and Lillian Mills. "The Evolving Schedule M-3: A New Era of Corporate Show and Tell?" *National Tax Journal* 57 No. 3 (September 2004): 757–72.
- Boynton, Charles, and William Wilson. "A Review of Schedule M-3, The Internal Revenue Service's New Book-Tax Reconciliation Tool." *Petroleum Accounting and Financial Management Journal* 25 No.1 (Spring

- 2006): 1–6. Available at <http://www.irs.gov/Businesses/Corporations/Published-Articles-on-Schedule-M-3-by-IRS-Treasury-Authors>.
- Internal Revenue Service. “Tax Information Security Guidelines for Federal, State, and Local Agencies and Entities.” *Publication 1075* (Rev 2007).
- Mills, Lillian. “Book-Tax Differences and Internal Revenue Service Adjustments.” *Journal of Accounting Research* 36 No. 2 (Autumn 1998): 343–56.
- Mills, Lillian, and George Plesko. “Bridging the Gap: A Proposal for More Informative Reconciling of Book and Tax Income.” *National Tax Journal* 56 No. 4 (December 2003): 865–93.
- Plesko, George A. “Reconciling Corporate Book and Tax Net Income, Tax Years 1996-1998.” *Statistics of Income Bulletin* 21 No. 4 (Spring 2002): 1–16.
- Plesko, George A., and Nina L. Shumofsky. “Reconciling Corporate Book and Tax Net Income, Tax Years 1995-2001.” Data Release. *Statistics of Income Bulletin* 24 No. 4 (Spring 2005): 103–108.
- Talisman, Jonathan. “Corporate Tax Shelters and the Corporate Tax Base,” pages 4–6, in “Penalty and Interest Provisions, Corporate Tax Shelters.” Testimony of Jonathan Talisman, Assistant Secretary (Tax Policy), U.S. Department of the Treasury, before the U.S. Senate, Committee on Finance, Washington, D.C., March 8, 2000.
- Towery, Erin. “The Interplay Between Mandatory Disclosure Regulation and Voluntary Disclosure: Evidence from Schedule UTP.” University of Georgia working paper, 2015.
- U.S. Department of the Treasury. “Evidence of Growth in Corporate Tax Shelters.” *The Problem of Corporate Tax Shelters: Discussion, Analysis, and Legislative Proposals*. Washington, D.C.: Government Printing Office, July 1999: 31–33.
- U.S. Office of Management and Budget. *Report on Statistical Disclosure Limitation Mythology*. Working Paper 22, (Rev. 2005).

Individual Nonfilers and IRS-Generated Tax Assessments: Revenue and Compliance Impacts of IRS Substitute Assessments When Taxpayers Don't File

Saurabh Datta, Stacy Orlett, and Alex Turk, Small Business/Self-Employed Division, Internal Revenue Service¹

Background and Introduction

The U.S. income tax system relies on taxpayers voluntarily filing tax returns when required, and reporting and paying their tax liabilities. Each year, a fraction of taxpayers fail to file required returns. After the filing season, the IRS identifies potential nonfilers and attempts to secure returns via a series of notices and other contacts. When the taxpayers fail to respond by filing a return, the IRS can file a “substitute for return” that creates a tax assessment based on prior-year information and information obtained from third parties. Many of the substitute assessments are made via the IRS’s Automated Substitute for Return (ASFR) process.²

The number of delinquent returns worked in the ASFR process varies from year to year. In recent years, the ASFR program has experienced a noticeable decline in resources and a corresponding decline in the number of delinquent returns worked by the ASFR process. This is partly attributable to a general decline in IRS budgets and partly due to the reallocation of nonfiler resources to other areas as IRS responsibilities expand. One criticism of the ASFR process is that in some cases the assessments are overstatements of the taxpayer’s true liability. Most of the deductions and exemptions a taxpayer may be entitled to claim can be obtained only if the taxpayer files the return. Thus, the ASFR assessments may overstate the true amount of unpaid tax. However, not making the ASFR assessments leads, in many cases, to an understatement of the unpaid tax that is owed to the U.S. Government. Another criticism is that many of the ASFR assessments can be difficult to collect.

To make the best use of the resources available to the IRS, it is critical that the IRS and policy makers understand the impacts of the ASFR program on collecting delinquent taxes and fostering future filing and payment compliance. To explore these impacts, we develop models of the potential collection of ASFR assessments and then predict the impact of the ASFR program on subsequent filing compliance. We can then use these models to estimate, in terms of dollars collected and the numbers of delinquent returns, the opportunity costs of reductions to the number of cases worked in the ASFR program.

In the next section, we broadly highlight some of the recent research papers in the area of taxpayer compliance and behavior, which will help us in understanding the noncompliance issues and then formulate the economic and empirical model from ASFR’s perspective.

Literature Review

The literature on taxpayer compliance is varied and has witnessed growth in recent years due to research in academics, public policy, and in legal enforcement.

¹ The views and opinions presented in this paper reflect those of the authors. They do not necessarily reflect the views or the official position of the Internal Revenue Service.

² ASFR is an automated process that generates notices and an automated “Substitute for Return” tax assessment if taxpayers do not resolve their delinquent returns. However, labor resources are needed to work with the taxpayer when the taxpayer responds to one of the letters or the default assessment.

Many economic models have been built to understand the interaction between taxpayers and tax authorities. The models developed in the literature are based on a principal-agent framework with highly simplified assumptions that fail to understand the complex relationship between the two parties. Understanding a taxpayer's behavior, psychology, moral, and social influences are critical elements in studying taxpayer compliance. Underestimating these factors has resulted in greatly overestimating noncompliance (Andreoni, *et al.* (1998)). Taxpayers' behavior may often appear to be "unethical," selfish, or irrational. Different taxpayers may behave differently under distinct circumstances, and some taxpayers may behave differently inter-temporally. However, taxpayers are not always driven by "unethical" traits, but are constrained by "bounded rationality" and underestimate the consequences of noncompliance (Alm and Torgler (2011)).

In recent years, attempts have been made to develop game-theory-based models intended to be more realistic, taking into account the repeated interactions between taxpayers and the tax authority. Considering the channels of interaction between taxpayers and the tax authority through notices and telephonic conversations may provide a more realistic model formulation and precise estimation (Andreoni, *et al.* (1998); Hashimzade, *et al.* (2013)). These models have enabled researchers to estimate both compliance and enforcement aspects simultaneously and address the problem of endogeneity in enforcement activities. However, on the empirical side, these models have broadly ignored some key aspects such as 100 percent document matching with third-party data (Plumley (1996); Kleven, *et al.* (2011)). The noncompliance rate and underreporting are high in cases of self-reported income, but in cases of income reported by third parties the tax evasion rate is very low. Therefore, supplementing tax administration data with household-level surveys and other governmental sources help in the detection of underreporting and improving voluntary compliance. Since data are collected for purposes other than for tax administration, the households and businesses are perhaps more likely to report their correct income and income sources. The magnitude of underreporting for self-employed businesses was estimated to be nearly 25 percent using U.S. data compared with tax administration data (Hurst, *et al.* (2014)). Moreover, most of the models assume audit rates to be fixed rather than endogenous. Audits as an endogenous tool of enforcement has resulted in greater compliance among self-reported taxpayers in a household study based in Denmark (Kleven, *et al.* (2011)).

The prime motivation of our research paper comes from Erard and Ho (2001), who specifically look at the issue of nonfilers with less restrictive assumptions. The authors incorporate nonfiling strategies adopted by nonfilers in a standard neoclassical theoretical model. The theoretical model accounts for sequential steps involved in a taxpayer's decision-making process that makes him decide whether to be compliant or non-compliant looking at the expected payoff from each decision point. Based on the theoretical model, the paper estimates a simultaneous equation model where simultaneity exists between the probability that a taxpayer will file a return and the likelihood of the taxpayer being located. The paper uses a 25 percent random sample of the IRS TCMP Phase III Survey, which has 54,000 individual income tax returns for Tax Year 1988. This research identifies key behavioral, demographic, and financial factors that influence a taxpayer's decision to file a tax return. A comparison between filers and nonfilers suggests that nonfilers have relatively fewer offsets and itemized deductions compared to the former. Unlike the income sources of the filers, a nonfiler's income mostly comprises business income and capital gain receipts. Moreover, the taxpayers who are close to the filing threshold seem to be deterrent to filing as the burden appears to outweigh the benefits.

Deterrence in noncompliance accrues when a nonfiler is treated directly by the tax authority for noncompliance in the form of audit or imposition of penalties and interests in addition to the unpaid tax liabilities. This is also known as the "direct" effect. The direct effect of enforcement has positive effect on voluntary compliance in the subsequent years for the treated nonfiler. Additionally, a change in compliance behavior may be triggered if the general population becomes aware of a change in enforcement level by the tax authority. This may result in deterrence in noncompliance. This effect is termed as "indirect" or "induced" effect in the existing literature (Plumley, 1996). Plumley (1996) has shown that the deterrence effect of audits on taxpayers is 11 times larger than the audits by themselves in his study involving data over a 10-year period. However, the estimation of "indirect" effect has been mainly confined to the tax audit literature. This concept has important economic and policy significance in the realm of filing and payment compliance and the ASFR program.

The aim of this research paper is not only to study the direct impact of ASFR on revenue collection and subsequent voluntary compliance, but also to estimate the indirect effect of this treatment on the general

nonfiler population. In doing so, the paper improves and extends the analysis of previous research in this area and contributes to the literature of taxpayer compliance by outlining a more realistic theoretical model considering imperfect information in a principal-agent framework. The empirical estimation from this research paper uses comprehensive taxpayer data, which are a substantial improvement over the predecessors who have been largely deficient in their analysis due to limited data availability.

Theoretical Model

We develop a theoretical model following tax compliance models proposed by Allingham and Sandmo (1972), and Andreoni, Erard, Feinstein (1998) using an expected utility maximizing framework within a principal-agent model setting—the principal being the tax authority and the agent being the taxpayer.

We assume that a representative (partly noncompliant) taxpayer who doesn't file an income tax return (nonfiler), but who has total income Y (some of which is unknown to IRS), has an amount θW of tax withheld from his "visible" income W (known to IRS). In the second step, only a certain portion of his income is reported on information returns (Y_r) while he suppresses the residual component of his income Y_s . However, the IRS, through its direct and indirect intervention, comes to know a proportion α of suppressed income Y_s . In the next step, if the nonfiler's reported income satisfies the criteria of the ASFR program, the case may be assigned to the ASFR treatment stream, and the IRS may identify another portion of the undisclosed income. Based on this logical structure, we propose to formulate a nonfiler's optimization problem and identify the instruments available to the IRS to promote taxpayer compliance.

Mathematically, we can formulate this problem as follows:

Let a nonfiler's total income equal:

$$Y = W + Y_s + Y_r$$

Where:

W = Income known to the IRS (on which tax is withheld);

Y_s = Income suppressed by the nonfiler; and

Y_r = Income reported for the nonfiler on information returns.

Assuming that α ($0 < \alpha < 1$) proportion of suppressed income Y_s is known to the IRS through its direct and indirect interventions,³ one can decompose his suppressed income level as $Y_s = \alpha Y_s + (1 - \alpha) Y_s$, which means

$$Y = W + \alpha Y_s + (1 - \alpha) Y_s + Y_r \text{ where } \alpha = \alpha(IRS) \text{ with } \alpha' > 0$$

In other words, $Y_r = Y - W - \alpha Y_s - (1 - \alpha) Y_s$

$Max [Y_r] = (Y - W)$ if $Y_s = 0$. This means the nonfiler's income is completely transparent to the IRS.

$Min [Y_r] = 0 \Rightarrow Y_s = (Y - W)$ This is a nonfiler who is suppressing all non-withheld income. Our presumption is that even then, the IRS can approximate the nonfiler's true income and attempt to extract the rightful taxes due through successive efforts.

Now, if p is the probability that a given nonfiler who has suppressed a part of his income is assigned to the ASFR treatment stream, then one can reasonably assume that:

p = Probability that the given nonfiler is selected to be treated in ASFR, where $p = p(ASFR)$ with $p' > 0$.

Due to ASFR intervention, the IRS may be able to get more information about the suppressed portion of the taxpayer's income. Let the proportion of suppressed income from Y_s as identified by ASFR be β , where:

³ Direct and indirect interventions include obtaining information about wages, income, interest, dividend, pension and social security incomes, etc. from direct and third-party reported sources.

$\beta = \beta(ASFR)$ with $\beta' > 0$ and $0 < \beta < 1$. This means at this stage that revelation of additional income through ASFR efforts is $\beta(1-\alpha)Y_s$.

Based on ASFR's intervention, the implied tax liability of a taxpayer would now be:

$$\text{Implied Tax Liability} = \theta(W+Y_r) + (\theta + \gamma)\alpha Y_s + T_1 + T_2 + (\rho + \theta)\beta(1-\alpha)Y_s$$

Here θ is the proportional tax rate, γ is the proportional penalty rate at the time of IRS assessment, ρ is the proportional penalty rate imposed by ASFR on the revealed portion of the suppressed income, T_1 and T_2 are flat-rate penalties assessed by the IRS in the beginning while matching direct and third-party reported documents and ASFR, respectively. For simplicity, assume this is a one-period static model, where everything is taking place within the same period, so there is no interest charge imposed on the taxpayer.

$\theta(W+Y_r)$ is the potential tax liability in the first stage on the withheld and revealed portion of the taxable income as reported for the nonfiler.

$(\theta+\gamma)\alpha Y_s + T_1$ is the tax and penalty charged by the IRS on the identified part of the suppressed portion of the income in the second stage (beyond the withholding stage).

$(\rho+\theta)\beta(1-\alpha)Y_s + T_2$ is the tax and penalty charged by the ASFR treatment stream on the further identified part of the suppressed income in the third stage.

Based on all this available information, a nonfiler's optimization problem can be stated as:

$$\begin{aligned} \underset{w.r.t. Y_s}{\text{Max}} EU(C) &= p(ASFR) \cdot U[Y - \{\theta(W+Y_r) + (\theta + \gamma)\alpha Y_s + T_1 + T_2 + (\rho + \theta)\beta(1-\alpha)Y_s\}] + \\ &\quad (1 - p(ASFR)) \cdot U[Y - \{\theta(W+Y_r) + (\theta + \gamma)\alpha Y_s + T_1\}] \\ &= p(ASFR) \cdot U[Y - T - \theta\{Y - Y_s(1-\alpha-\beta(1-\alpha))\} - Y_s\{\gamma\alpha + \rho\beta(1-\alpha)\}] + \\ &\quad (1 - p(ASFR)) \cdot U[Y - T_1 - \theta\{Y - (1-\alpha)Y_s\} - \gamma\alpha Y_s] \end{aligned}$$

$$\text{Let } A = Y - T - \theta\{Y - Y_s(1-\alpha-\beta(1-\alpha))\} - Y_s\{\gamma\alpha + \rho\beta(1-\alpha)\}$$

$$B = Y - T_1 - \theta\{Y - (1-\alpha)Y_s\} - \gamma\alpha Y_s$$

We assume $T = T_1 + T_2$. An underlying assumption is that the tax liability of the nonfiler is greater than their withheld amount. That is,

$$E(\text{Tax liability with penalties and interest}) \geq \theta W$$

Also, for simplicity, assume that a nonfiler has a Constant Relative Risk Aversion (CRRA) utility function (Ljungqvist and Sargent, 2000):

$$U(C) = \frac{C^{1-\mu}}{1-\mu}$$

where $0 < \mu < 1$, and μ signifies risk aversion, with $U'(C) > 0$ and $U''(C) < 0$.

Based on the choice of utility function, the first order conditions can be derived as:

First Order Condition w.r.t. Y_s :

$$p(ASFR) \cdot A^{-\mu}[\{\theta(1-\alpha-\beta(1-\alpha))\} - \{\gamma\alpha + \rho\beta(1-\alpha)\}] + (1 - p(ASFR)) \cdot B^{-\mu}[\theta(1-\alpha) - \gamma\alpha] = 0$$

The first order condition needs to be solved for Y_s given the values of ρ , θ , α , β , p , T_1 , T_2 , and Y .

Y_s will be a nonlinear equation expressed in terms of the exogenous parameters and the instruments under IRS-ASFR's control, which can be solved using numerical simulations.

$$\text{Let } C = \{\theta(1-\alpha-\beta(1-\alpha))\} - \{\gamma\alpha + \rho\beta(1-\alpha)\}$$

$$D = \theta(1-\alpha) - \gamma\alpha$$

Second Order Condition w.r.t. Y_s can be written as:

$$p(ASFR) \cdot \mu A^{-\mu-1} \cdot C^2 - (1 - p(ASFR)) \cdot \mu B^{-\mu-1} \cdot D^2 < 0$$

$$\Rightarrow -[p(ASFR) \cdot \mu A^{-\mu-1} \cdot C^2 + (1 - p(ASFR)) \cdot \mu B^{-\mu-1} \cdot D^2] < 0$$

The above expression is true since all the elements within the parentheses are positive quantities by construction.

In this model, there are two special cases, which deserve attention:

1. A solution value of $Y_s = 0$ suggests the IRS is in a position to set up parameter values that produce 100-percent compliance.
2. $Y_s = (Y - W)$ suggests the IRS has no income information beyond W ; that is $Y_r = 0$. This is a case of a nonfiler. The IRS has to work hard through several successive steps to obtain an estimate of Y and impose tax on the nonfiler.

The objective of the IRS and ASFR may be viewed as:

Min $Y_s = Y_s(\rho, \theta, \alpha, \beta, p, T_p, T_2)$ subject to the resource constraints and based on available information on Y . A combination of the parameters $\rho, \theta, \alpha, \beta, p, T_p, T_2$ and a realistic estimate of Y will minimize the value of Y_s .

Interestingly, this model captures the natural conflicts of interest—while a taxpayer may be interested in maximizing his utility through choice of suitable value of Y_s , the IRS would like to see a minimum, even a zero value of Y_s , and thus would like to choose an appropriate combination of values of the parameters under its control. The present model formulation does not explicitly cover the process through which the IRS chooses the right combination of values of these parameters (i.e., how the IRS chooses Y_s as a function of these parameters from the taxpayer's reaction function as derived above). The final values of Y_s and the underlying parameters, which may be interpreted as Nash equilibrium, will invariably be the result of a convergent sequence of interactions (provided it exists) between the two sides over time (once we introduce explicit and discrete time lags based on observed behavior). However, the IRS never operates with unlimited budget to choose any combination of values of the underlying parameters (including a predicted value of Y). So, more often than not, the model formulation from the IRS side must explicitly incorporate how over time the IRS manages to ease its budget constraint, besides undertaking certain necessary reforms to improve effectiveness of its systems and process.

However, an achieved Nash equilibrium is still not a Pareto-optimal move. In order to achieve Pareto optimality or Nash bargaining equilibrium, the IRS and the taxpayer must be engaged in a 'Coasian' negotiation process (Coase, 1960; Milgrom and Roberts, 1992).

The above stated theoretical model suggests the purpose of the ASFR program is to minimize Y_s and thereby maximize tax revenue (or Dollars Collected) and then promote voluntary subsequent compliance in the successive years. The factors that minimize Y_s help in identifying the taxpayer's characteristics that maximize the expected revenue for the IRS based on this principal-agent framework.

Now, suppose ASFR treatment is applied to n_1 taxpayers out of n , where n were eligible for ASFR treatment.⁴ In the first stage of the estimation process, we estimate the probability of a taxpayer's case being selected for ASFR treatment (p). The probability can be estimated based on observable taxpayer characteristics. In the second stage, given the probability that ASFR has worked the case, we estimate the revenue collection over a period of the next 3 years.⁵ We assume that the dollars collected over the next 3 years depends on ASFR treatment, the probability of being worked by the ASFR treatment stream, and other observable taxpayer characteristics beyond those captured in the probability measure. In other words, there are certain factors that affect

⁴ Since ASFR doesn't have unlimited resources, it selects to work certain cases based on the observable characteristics of the taxpayer.

⁵ Three years is selected based on the maximum time an ASFR case generally takes time to resolve under normal circumstances.

revenue collection that do not affect the probability of being selected in ASFR, which are suitably introduced in this model specification.

The variable signifying ASFR treatment captures the direct treatment effect of ASFR on revenue collection, whereas the probability measure estimates the indirect effect of ASFR. The indirect effect captures the additional productivity of the case due to the reason that it has been worked by ASFR. This measure provides an empirical evidence of ASFR's additional effectiveness in collecting revenue by employing its collection instruments beyond its direct effects.

Furthermore, we estimate the future voluntary compliance of taxpayers who have been through the ASFR treatment process earlier. We specifically estimate the voluntary compliance of these taxpayers 2, 3, and 4 tax years later.⁶ The factors affecting future voluntary compliance depend on a taxpayer's past observable characteristics, previous treatment by ASFR, probability that the case was assigned earlier to ASFR, and whether the taxpayer self-corrected before the next tax return was due. We argue that the taxpayer who self-corrected before the next tax return was due demonstrated greater willingness to be compliant and therefore needs to be controlled for suitably in the empirical model. Analogous to the earlier model, the indicator that ASFR worked this taxpayer's case previously and the probability of being worked by ASFR earlier measures the direct and indirect effects of ASFR on taxpayer's future voluntary compliance respectively.

In the following sections, we explain in detail the delinquent process, explain the data sources and variables selected, and then estimate the empirical models.

Summary of the IRS Return Delinquency Process

The IRS individual Return Delinquency process is illustrated in Figure 1. The process begins by identifying individual taxpayers who may be required to file but have not filed a tax return (e.g., Form 1040) by the Return Due Date.⁷ A primary method the IRS uses to identify these taxpayers is the individual Return Delinquency case creation process. The case creation process is critical for identifying income and other information for these taxpayers, which is reported by third parties to the IRS on several types of information returns (e.g., Form W-2, Form 1099-R, etc.).⁸ Using this information alone will not identify all nonfilers, but it will identify those who have some sort of income that is reported to the IRS. We can refer to these as the *known* nonfilers. An example of an *unknown* nonfiler would be an individual who has only cash income that is not reported on information returns and therefore is not identified during the case creation process.

After some additional compliance checks, a portion of these known nonfilers will be identified as required to file and will go into the Return Delinquency notice process. The nonfilers entering the notice process will receive up to two notices requesting them to file their tax return. During the notice process, a taxpayer has up to 14 weeks to respond. If a taxpayer does not respond to these notices, the case may then proceed to Taxpayer Delinquent Investigation (TDI) status.

The types of treatment a nonfiler in a TDI status receives varies based on case characteristics. A TDI may be treated by various functions including the Automated Collection System or Call Site (ACS), a Revenue Officer from a Field Collection office (FC), the ASFR program, and/or others. The ASFR program is an important program to the IRS for enforcing filing compliance by determining and assessing a tax liability when the taxpayer has not come forward with a return.⁹ Some TDI cases go directly to the ASFR inventory after the notice process and some of the ASFR inventory is transferred to the ASFR function from other functions (e.g., ACS or FC) after unsuccessful attempts to secure or otherwise resolve the delinquent return, as long as the case meets specific ASFR eligibility criteria.¹⁰

⁶ The very next tax year may be too short of a period of time to measure voluntary compliance. Therefore we selected 2 tax years later. Also, the results for 3 tax years later are consistent with the 2 years later results.

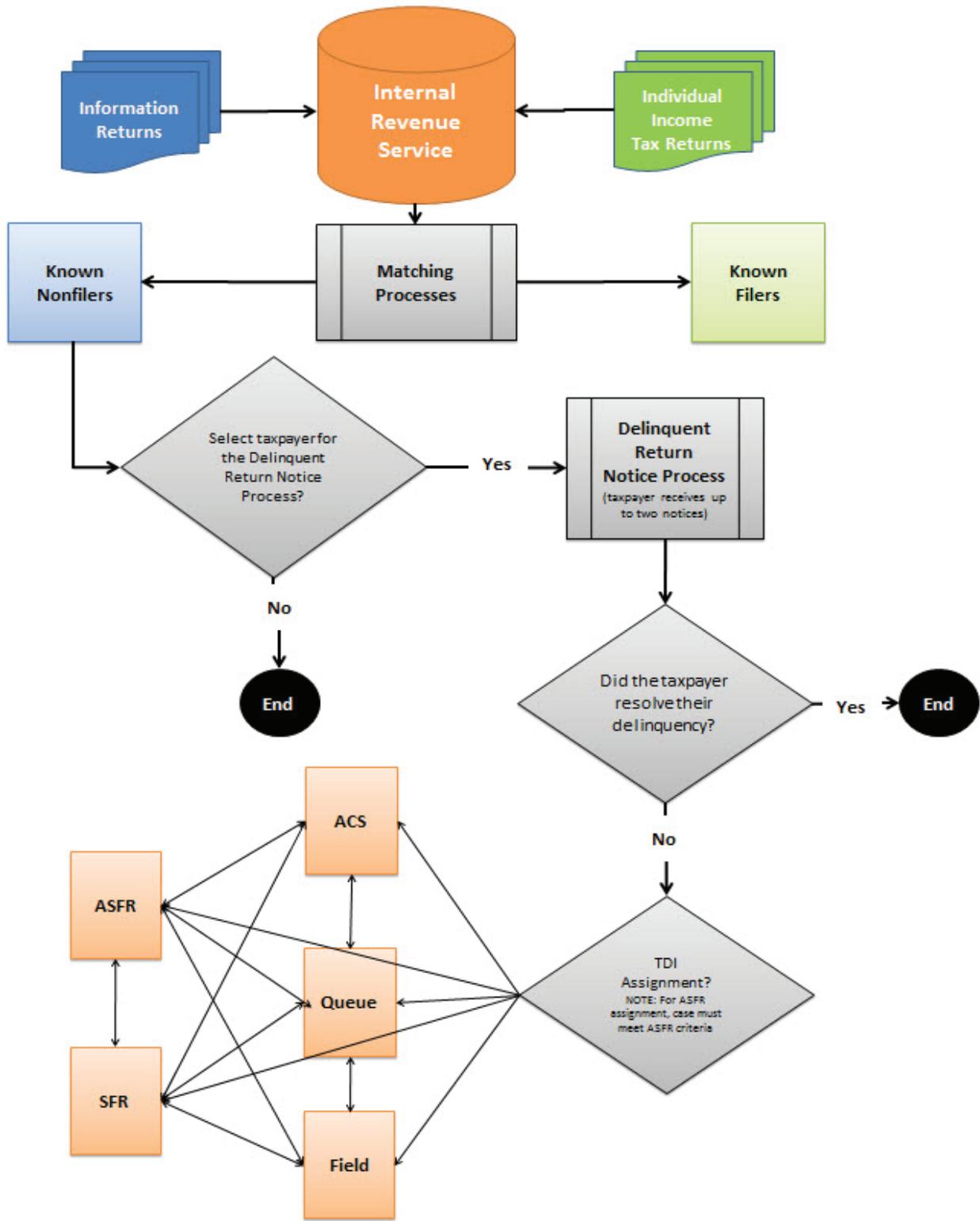
⁷ Internal Revenue Service. Internal Revenue Manual 5.19.2.1 (01-16-2015) "What is the IMF Return Delinquency Program?"

⁸ Internal Revenue Service. Internal Revenue Manual 5.19.2.4.1 (01-16-2015) "IRP Income."

⁹ Taxpayers must file a return if they wish to demonstrate that their tax liability is different than the ASFR assessment. Payments of assessed amounts can be made by the taxpayer, offset from refunds claimed by the taxpayer for other tax years, or generated from liens on the taxpayer's assets or levies on the taxpayer's income sources.

¹⁰ Internal Revenue Service. Internal Revenue Manual 5.18.1.3.1 (12-09-2014) "ASFR Criteria."

FIGURE 1. IRS Return Delinquency Process



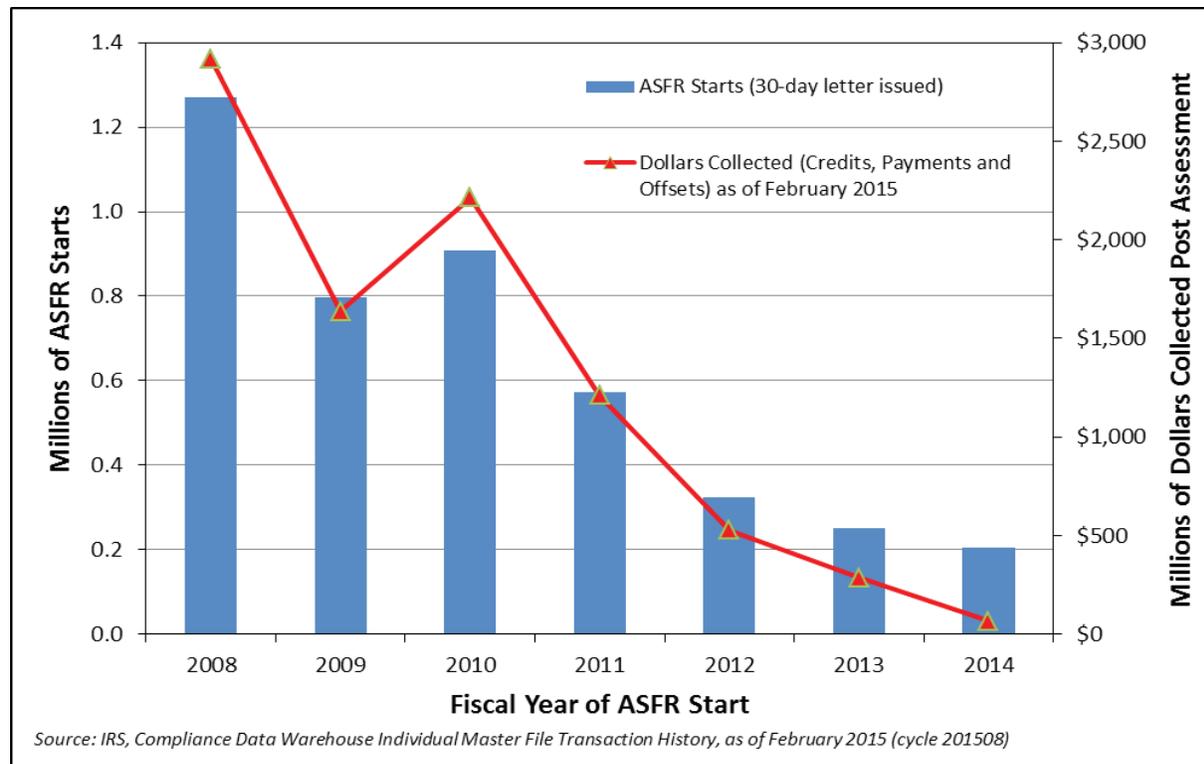
Source: Internal Revenue Service, Internal Revenue Manual 5.19.2.1

Cases received by the ASFR program are prioritized by Refund Hold, Tax Year, and Net Tax Due.¹¹ For purposes of our research, we are excluding Refund Hold cases worked by ASFR to focus on the discretionary TDI cases worked and available to be worked. Refund Hold cases must be worked within a designated amount of time and differ from non-Refund Hold because the service is holding a refund for that taxpayer from another tax return. The taxpayer will be notified that the Service is holding the refund, and that he must resolve all of his delinquent returns within the last 5 years prior to the current year before the IRS releases the refund. By holding the refund, the taxpayer, arguably, has different motivations to file the delinquent tax returns compared to the non-Refund Hold cases.

The ASFR process begins with various compliance checks. If the taxpayer passes these checks, a 30-day letter (Letter 2566) will be systemically sent to the taxpayer giving him 30 days to respond. If there is not a sufficient response from the taxpayer to the 30-day letter, then ASFR generates a Statutory Notice of Deficiency (also known as the 90-day letter) sent by certified mail. If there is still not a sufficient response from the taxpayer after the 90-day letter is issued, then the ASFR process will systemically request a default assessment based on the proposed tax assessment.

Figure 2 provides the number of cases started by ASFR, where a 30-day letter was issued, over the last 7 fiscal years. Historically, ASFR has had more cases available to initiate than available resources to work them. As ASFR resources have declined over the years, fewer cases have been started. In FY 2008, over 1.2 million cases were started by ASFR compared to only approximately 200,000 in FY 2014. The line on the figure represents the dollars collected to date for respective ASFR starts in that fiscal year. Since the cases started in FY 2008 have had more time to collect compared to cases worked in FY 2014, the decline in dollars collected to date is due to both the decline in the number of cases and the shorter time (so far) to collect. On cases ASFR started in FY 2008, the IRS has collected over \$2.5 billion to date. The decline in cases worked also means fewer returns secured and lower dollar amounts collected on these delinquent returns.

FIGURE 2. ASFR Starts (30-day letters issued) During Fiscal Years 2008-2014



¹¹ Internal Revenue Service. Internal Revenue Manual 5.18.1.3.2 (06-20-2012) "ASFR Prioritization."

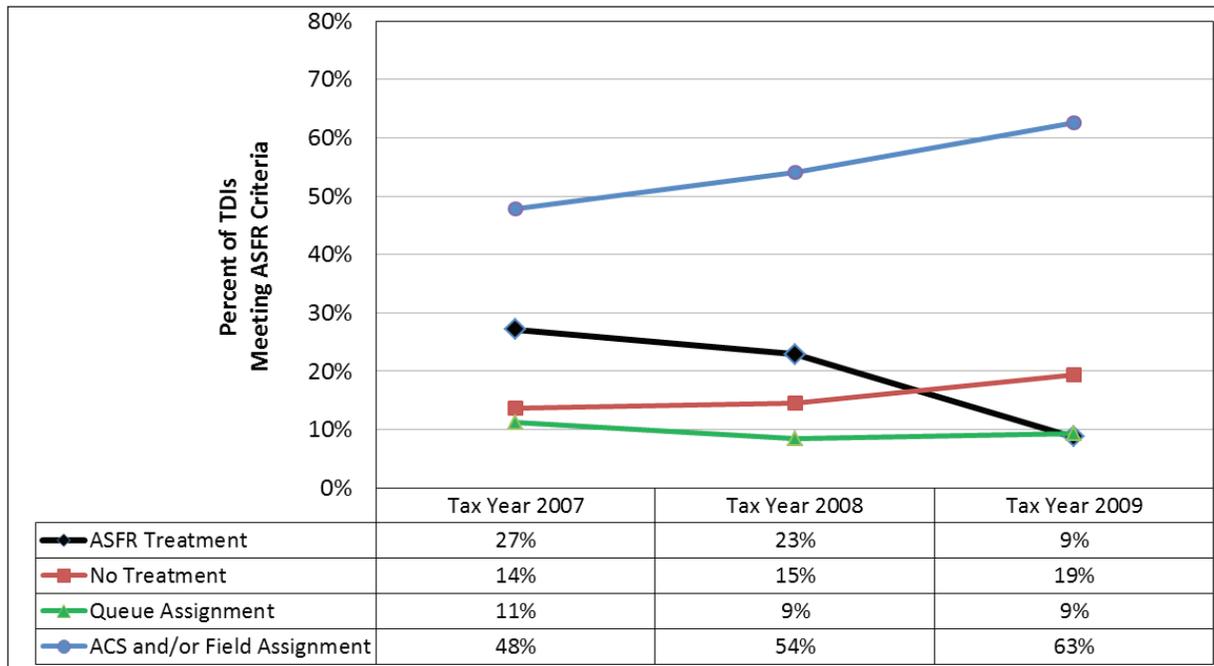
Research Design

Available ASFR Inventory

To develop our models, we identified a set of cases that would be available for ASFR to work. To begin, we identified individual delinquent tax returns for Tax Years 2007, 2008, and 2009 that eventually became TDIs.¹² Using these tax years allows us to capture various changes from year to year, and enough time to evaluate the compliance behavior from TDI status in terms of dollars collected and subsequent filing compliance. Next, we identified the TDIs that met the ASFR criteria.

Figure 3 provides the number of TDIs for the 3 years that met ASFR criteria and the types of treatments received, if any. Of the cases that met ASFR criteria, the percent of inventory being treated by ASFR declined from TY 2007 to TY 2009. As ASFR treats fewer cases, the percent of cases not treated or assigned to ACS and/or Field increased.

FIGURE 3. Treatments and Assignments of Taxpayer Delinquent Investigations Meeting ASFR Criteria, Tax Years 2007-2009

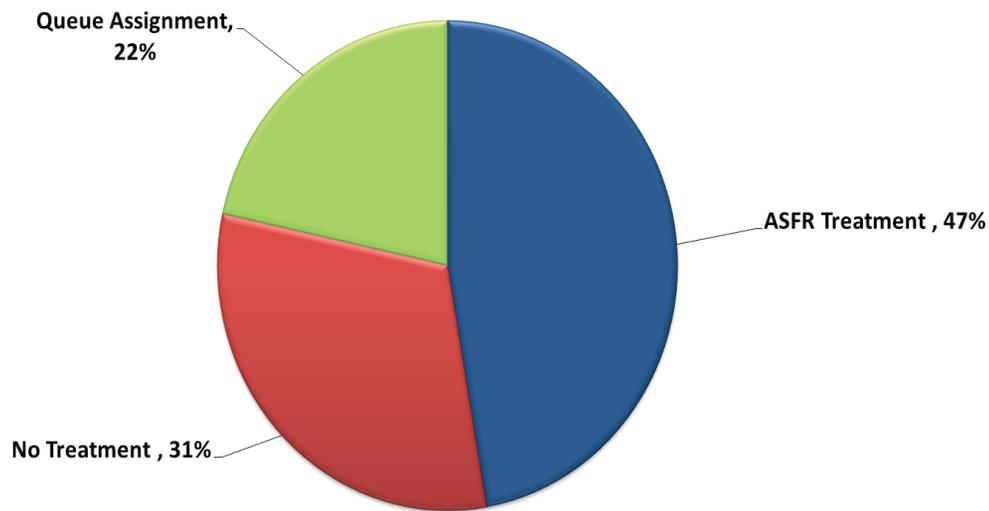


For this paper, we analyzed the TDIs that met ASFR criteria and were available for ASFR to work. TDIs available for ASFR to work were identified as those not assigned to ACS or Field functions. This left us with a set of TDIs available for ASFR that were either treated by ASFR, given no treatment, or assigned to the Queue.¹³ Figure 4 provides the percentage of the available TDIs that fall into each of these three categories for Tax Years 2007–2009. TDIs treated by ASFR accounted for 47 percent of the TDIs available for ASFR over these 3 tax years.

¹² ASFR inventory includes different types of cases that are classified into distinct priority categories based on their observable characteristics. Each year a stratified sample representing each priority category is selected to be worked from this inventory. Furthermore, the ASFR resources have varied widely across years and thus there has been a large variation in the types of cases selected in the ASFR program each year. We use this randomized representation of ASFR inventory to identify ASFR treatment effects.

¹³ While these TDI Cases are not assigned randomly, the only factors that influence the likelihood of a case being worked in the ASFR process is a narrow set of case characteristics and available resources. Available resources have varied widely across the years included in this study and are unrelated to taxpayer behavior.

FIGURE 4. Taxpayer Delinquent Investigations Meeting ASFR Criteria and Available for ASFR To Work, Tax Years 2007–2009



Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

We used several IRS databases to identify the TDIs available for ASFR to work.¹⁴ All results in this paper are based on a 10 percent random sample.¹⁵

Dependent Variables

This research focuses on compliance behavior in terms of both payment and filing compliance.

1. **Dollars Collected.** We modeled the dollars collected related to the delinquent return. We aggregated, to the case level, the dollars collected over the 3 years following TDI status. This provided a consistent time frame for everyone in our research design.
2. **Subsequent Filing Compliance.** We define subsequent voluntary filing compliance as a dichotomous outcome. We assigned a “1” if the taxpayer voluntarily filed a subsequent return and a “0” if not. A voluntarily filed return is defined as one filed by the return’s due date (e.g., April 15), or by the requested extension date, without any subsequent delinquent return treatments. To be consistent across all cases in our study, we noted whether the taxpayer voluntarily filed the tax return associated with tax periods 2, 3, or 4 tax years following the delinquent return tax period. For example, for a taxpayer with a TY 2007 delinquent return in our study, we identified if they voluntarily filed a TY 2009, TY 2010, or TY 2011 returns.¹⁶

For dollars collected, 28 percent of the TDIs treated by ASFR made a payment within 3 years of becoming a TDI. On average, \$1,454 dollars were collected per case treated by ASFR. Cases without treatment or assigned to the Queue had a lower percentage of taxpayers with a payment and on average fewer dollars collected compared to cases treated by ASFR. The table below provides a summary of the dollars collected on the TDIs available for ASFR to work for Tax Years 2007–2009.

¹⁴ Data was gathered from Individual Case Creation Nonfiler Identification Process, Individual Master File Status and Transaction History, and Individual Return Transaction File databases stored on the IRS Compliance Data Warehouse.

¹⁵ The ASFR starts in our modeling population that had been previously assigned to other treatment streams, namely Automated Collection System (ACS) and/or Collection Field Function (Cf) accounted for 26 percent. These are indirect assignments to ASFR. This characteristic is suitably controlled in the empirical model.

¹⁶ We did not control for taxpayers not having a filing requirement in the subsequent voluntary filing compliance models. Taxpayers may have had economic circumstances removing their filing requirement.

TABLE 1. Dollars Collected Within Three Years from TDI Status on TDIs Available for ASFR (Tax Years 2007–2009)

Type of Treatment	Percent with a Payment	Average Dollars Collected (all cases)	Cases with a Payment		
			25th Percentile Dollars Collected	50 th Percentile Dollars Collected	75 th Percentile Dollars Collected
ASFR Treatment	28%	\$1,454	\$805	\$2,147	\$4,914
No Treatment	19%	\$804	\$723	\$1,975	\$4,150
Queue Assignment	6%	\$384	\$491	\$1,557	\$4,996

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

Note that cases were not randomly assigned to these three treatments, so the differences cannot be attributed solely to the treatments.

For subsequent compliance, we chose to look at the second to the fourth years following the delinquent return. This time frame was chosen based on the amount of time it takes the IRS to identify a delinquent return and go through the delinquent return process before becoming a TDI. Take, for example, a TY 2007 return with a return due date of April 15, 2008, for which the taxpayer filed for an extension to October 15, 2008. During this time, the IRS received information returns from third parties providing income and other information about taxpayers. Once any granted extensions have passed, the IRS begins the case creation process of identifying delinquent returns that appear required to be filed based on income reported on those information returns. From this process, the IRS can identify a set of potential nonfilers and select cases to put into the delinquent return notice process. Recall, this notice process can take approximately 14 weeks if the taxpayer does not respond. After that, a portion of these cases will then become TDIs. For a delinquent TY 2007 return, in general, the notice process was ongoing at the time the TY 2008 return was due. Therefore, we structured our study to look at the next TY return due, which was TY 2009, to allow time for the delinquent return process to begin. In addition, we wanted to look at the taxpayer's filing compliance for the subsequent 2 years: TY 2010 and TY 2011.

When analyzing compliance, a taxpayer's past compliance behavior can be a good predictor of his future compliance behavior. Therefore, we looked at how many of the taxpayers in our study had later filed their delinquent return prior to the due date of subsequent tax returns. Of the available TDIs not treated by ASFR, a portion eventually filed their return without any treatment. The table below provides the percentage of cases that filed the delinquent return prior to the second, third or fourth subsequent tax years. For example, of the cases with no treatment, at least 20 percent had filed their delinquent return prior to the due date of the third and fourth subsequent tax years.

TABLE 2. Return Filed After TDI Status on TDIs Available for ASFR Not Treated (Tax Years 2007–2009)

Type of Treatments for TDI Before Due Date of Subsequent Return	Percent of Cases That Filed Return on TDI Before Due Date of Subsequent Return		
	Two Tax Years After TDI	Three Tax Years After TDI	Four Tax Years After TDI
No Treatment	10%	20%	23%
Queue Assignment	4%	7%	8%

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

Note that cases were not randomly assigned to these three treatments, so the differences cannot be attributed solely to the treatments.

Next, we identified subsequent voluntary filing compliance using the same groups identified in Table 2. Table 3 provides the percentage within each group that voluntarily filed a subsequent return. We found that taxpayers who eventually filed their delinquent return for the TDI before the due date of the subsequent return

had a higher rate of voluntary subsequent filing compliance compared to taxpayers treated by ASFR or taxpayers not filing their delinquent return. Of the cases ASFR treated, 39 percent voluntarily filed a subsequent tax return 4 years after the tax year of the TDI.

TABLE 3. Subsequent Voluntary Filing Compliance on TDIs Available for ASFR (Tax Years 2007–2009)

Type of Treatments for TDI Before Due Date of Subsequent Return	Filed Return of the TDI Before Due Date of Subsequent Return	Percent of Cases That Voluntarily Filed a Subsequent Return		
		Two Tax Years After TDI	Three Tax Years After TDI	Four Tax Years After TDI
ASFR Treatment	No	33%	38%	39%
No Treatment	No	37%	42%	42%
	Yes	71%	69%	61%
Queue Assignment	No	22%	25%	26%
	Yes	65%	63%	56%

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

Note that cases were not randomly assigned to these three treatments, so the differences cannot be attributed solely to the treatments.

Independent Variables

Independent variables for our models included a dummy and other controls for ASFR treatment, and a variety of explanatory variables about the taxpayer and delinquent return gathered at the time the delinquent return case was created. To create a nonfiler case, the IRS identifies income and payments from various types of information returns, plus information reported and compliance characteristics on prior tax returns, and then computes a potential balance due.¹⁷

For modeling both dollars collected and subsequent compliance, we used the dummy for ASFR treatment to capture the average direct effect. To identify indirect effects, we created an instrumental variable to control for the probability the case is worked by ASFR. The instrumental variable was developed using a probit regression with the dependent variable being ASFR treatment. We used tax year as the instrument.

We also included in all the models a variable to control for the number of cycles from TDI status to ASFR treatment.

Empirical Model

We estimate three sets of regressions to assess the impact of ASFR on dollars collected and subsequent filing compliance. First, we estimate the probability of ASFR working a case from available ASFR inventory. Second, we estimate the dollars collected within the next 3 years following TDI status using a linear specification and a Tobit specification.

We then estimate a third set of regressions to assess the impact of ASFR on voluntary subsequent filing compliance 2, 3, and 4 years after the delinquent tax year of the TDI.

Model: Probability of ASFR Selection

Since the cases were not assigned to the groups randomly, the first regression provides an estimate of the probability that a case is selected for ASFR treatment using case characteristics and proxies for ASFR resources/level of treatment. The specification of the regression is as follows:

¹⁷ The potential tax liability is based on the assumption that the taxpayer is single with no itemized deductions, adjustments, or credits.

$$\text{Model 1: } P(\text{ASFR}=1)_i = \Phi(X_{im}, \beta_m)$$

In Model 1, the probability of ASFR working a case i is estimated as a function of a set of m exogenous variables, including a taxpayer's observable characteristics obtained from the case creation process, represented as X_{im} . The observable characteristics include various income and income sources of the taxpayer, past tax compliance behavior, and the nature of the taxpayer (Federal employee, small business, etc.), etc. In addition, there are dummies for the tax year of the TDI to serve as proxies for ASFR resources/level of treatment. Recall, this paper looks over multiple tax years, and during the time these tax years were available to treat the resources and levels of ASFR treatment changed as shown in Figure 2. As mentioned in the literature review, prior research has explored changes in compliance behavior as a result of changes in enforcement levels by the tax authority (Plumley (1996); Bloomquist (2004)). By computing the probability of ASFR working a case, regardless of treatment, we are capturing the indirect effect of changes in ASFR's level of enforcement and how that impacts a taxpayer's subsequent compliance behavior perhaps based on the taxpayer's perception of potential ASFR treatment.¹⁸ For example, as a result of the probability or likelihood of ASFR treating a taxpayer who hasn't filed their return declines does the taxpayer's future compliance behavior also decline?

The β coefficients estimate the impact of changes in X on the probability. We obtain these predicted probabilities using a probit regression.

Model: Dollars Collected

We estimate two specifications for the dollars collected. The first, a linear model of dollars collected, is:

$$\text{Model 2A: } Y_i = \beta_1 \text{ASFR}_i + \beta_2 P(\text{ASFR})_i + X_{ik} \beta_k + e_i$$

Y_i indicates dollars collected within 3 years of TDI status for case i , ASFR is an indicator for ASFR treatment, and $P(\text{ASFR})$ is the predicted probability from Model 1. X_{ik} is a vector of additional k observable taxpayer characteristics over and above the characteristics captured by X_{im} . The model is estimated using Ordinary Least Squares (OLS) regression. We tested for heteroscedasticity using White's Heteroscedasticity test (White, 1980) and used Heteroscedasticity Consistent Covariance Matrix to correct the standard errors. Since this is an OLS equation and the dependent variable is dollars collected, β_1 and β_2 capture the marginal direct and indirect effects of ASFR treatment on dollars collected from delinquent returns in the available inventory.¹⁹

As shown in Table 1, there are a number of cases in our sample for which the IRS has not received any payments from the delinquent taxpayer once they are in TDI status. This results in left censoring of the payment variable at zero dollars collected. In such a situation, the OLS estimates are inconsistent, the slope is biased upward, and the intercept is biased downward. A Tobit estimate using maximum likelihood estimation is consistent (Amemiya (1973)). Hence, we estimate another variant of Model 2A below using a Tobit regression:

$$\text{Model 2B: } Y_i = \beta_1 \text{ASFR}_i + \beta_2 P(\text{ASFR})_i + X_{ik} \beta_k + e_i$$

Where Y_i is a latent variable: $Y_i^* = 0$ if $Y_i \leq 0$ and $Y_i^* = Y_i$ if $Y_i > 0$. We use the same exogenous variables as specified in Model 2A. The parameters β_i reflect the marginal impacts of each variable on the latent variable, Y_i^* . The marginal impact on dollars collected is given by:

$$\frac{\partial(Y_i^*)}{\partial x_i} = \beta_i \Phi \left(\frac{(\beta_1 \text{ASFR}_i + \beta_2 P(\text{ASFR})_i + X_{ik} \beta_k)}{\sigma_U} \right)$$

Where x_i is a specific element of the set $[\text{ASFR}_i, P(\text{ASFR})_i, X_{ik}]$, $\Phi()$ is the normal distribution function and σ_U is the scale parameter.

¹⁸ Our model assumes taxpayers do not update their expectation of the probability they are selected as new tax years come due, especially when the taxpayer has not yet been selected.

¹⁹ We have controlled for the characteristics used for case selection. We assume there are no other idiosyncratic factors associated with the taxpayer that influence the likelihood of selection. Thus, the random component in the selection process is not related to the error in Model 2 or Model 3. If the random component in the selection process and the errors were in fact correlated with taxpayer behavior, we would need to use the probability of treatment as an instrumental variable for the direct treatment effect.

Model: Subsequent Filing Compliance

The third set of regression models estimate the impact of both direct and indirect effects of ASFR on subsequent voluntary filing compliance for ASFR inventory in subsequent years. We model subsequent voluntary filing compliance for tax period $t+j$ as:

$$\text{Model 3: } P(\text{File}_{t+j})_i = F(\alpha_1 \text{ASFR}_i + \alpha_2 P(\text{ASFR})_i + X_{ij} \alpha_j).$$

The variable File_{t+j} represents whether the taxpayer timely filed their $t+j$ tax return, X_{ij} represents case characteristics at time that return is due, and $F()$ is a logistic probability distribution function.²⁰ Model 3 provides estimates of ASFR direct and indirect effects on subsequent compliance j tax years after a delinquent return. In our case, $j = 2, 3,$ and 4 .²¹ We estimate separate models for each j . This model is estimated using a logistic regression. Since the parameter estimates obtained from a logistic regression are obtained in the form of log-odds ratio, we compute marginal effects of ASFR treatment and the probability of ASFR working a delinquent case on subsequent compliance. These marginal effects are important in testing our hypothesis that ASFR treatment has both direct and indirect effects on subsequent voluntary filing compliance.

Results

In this section we report our regression results of ASFR treatment on dollars collected and on subsequent voluntary filing compliance. Our sample consists of cases that satisfy the requirements of ASFR treatment and were available for ASFR to treat. However, due to resource constraints and prioritization within ASFR, only a portion of those identified cases are worked in ASFR.²² In the first step, we estimate the probability of ASFR working a case from its available inventory. The probability is estimated based on available observable characteristics of the taxpayer. The results are reported in the Appendix.²³

Table 4 reports the regression results for the dollars collected model for specifications 2A and 2B. The dependent variable for these regressions is defined as the net dollars and offsets²⁴ collected on a module for the 3 years starting when the delinquent return enters TDI status. The direct impact of ASFR treatment on dollars collected is captured by the indicator that the module has been worked by ASFR. The coefficient from the OLS²⁵ model and the marginal effect from Tobit models suggest that a case treated by ASFR, on average, yields \$672 and \$1,640 more, respectively, than a module that is not treated by ASFR, while keeping other factors fixed.²⁶ Recall the coefficient on probability of ASFR working a module captures the indirect effects of ASFR. In this case, the coefficient suggests that treating an additional case in ASFR provides an indirect increase in dollars collected for the OLS and Tobit model of \$194 and \$1,187 respectively,²⁷ all else equal. These two coefficients provide an estimate of both direct and indirect effects of ASFR working a module based on the representative sample we have used to estimate our models. For dollars collected, the direct effect of ASFR is larger than the indirect effects. Also the Tobit estimates are much larger than the OLS estimates. These marginal effects will be explored later via a simulation of working additional cases for Tax Year 2009.

²⁰ There is potential for future related research in subsequent voluntary filing compliance to control for taxpayers having a filing requirement based on the information reported for the taxpayer. Our definition of subsequent voluntary filing compliance does not account for when a taxpayer is not required to file subsequent tax years.

²¹ We do not select $j=1$ due to the lag in the ASFR treatment and the impact of ASFR treatment on subsequent compliance.

²² While there is some prioritization within ASFR, the only factors that influence the likelihood of a case being worked in the ASFR process is a narrow set of case characteristics and available resources. Available resources have varied widely across the years included in this study and are unrelated to taxpayer behavior.

²³ We consider Fiscal Year dummies instead of Tax Year dummies as instruments in this regression, but the results are insensitive to this choice.

²⁴ The net dollars and offsets includes any applied payments and offsets minus any reversed payments or offsets; such as misapplied payments, bad checks, etc.

²⁵ OLS estimates are marginal effects.

²⁶ We provide both OLS and Tobit model results to show the effect of when you do or do not control for the censoring.

²⁷ The indirect effect for a specific case is $\beta_2 * 1/N$. Since the indirect effect impacts all cases, the aggregate indirect effect of an additional case is β_2 .

TABLE 4. Models 2A & 2B: Expected Dollars Collected Three Years from TDI Assignment

Dependent Variable: Dollars Collected Three Years from TDI Assignment Explanatory Variables	Model 2A: OLS Estimates	Model 2B: Tobit	
		Estimates	Marginal Effect
Indicator of ASFR Treatment (ASFR)	\$672.44 (14.60)***	\$11,385.00 (182.15)***	\$1,639.59
Predicted Probability of ASFR Working a Case (P(ASFR))	\$193.92 (37.32)***	\$8,241.28 (340.38)***	\$1,186.86
Number of Cycles to ASFR Treatment (30-day letter issued)	-\$6.55 (0.24)***	-\$103.76 (3.26)***	-\$14.94

Source: Internal Revenue Service Individual Master File Status and Transaction History, and Case Creation Nonfiler Identification Process. Data extracted February 2015. Notes: Not all explanatory variables shown. See Appendix.

N = 277,314

*p<0.1; **p<0.05; ***p<0.01; Standard errors reported in parentheses; the standard errors for the OLS model are Heteroscedasticity Consistent Standard Errors. Marginal Effects are calculated at the sample means.

The estimated parameter for the number of cycles to the 30-day letter (the start of the ASFR process) is negative and significant in both models. One interpretation of this may be that the sooner the ASFR process is started after TDI status, the greater the amount of unpaid tax collected within the 3-year time frame. Thus, starting the ASFR process early will result in more dollars collected. While it seems intuitive that this would be the case, it may not be the appropriate conclusion with the model as specified. Due to data limitations, we cannot track all ASFR cases for a consistent amount of time between the ASFR case closing and any unpaid tax being assessed. Thus, cases that are started later have a shorter window of time for us to observe payments following ASFR treatment. The “cycle-to-30-day letter” measure is controlling for the varied amount of time following treatment to observe payments on each case worked by ASFR. The other control variables in the regression are significant. They are reported in the Appendix.

The individual t-statistics of each parameter estimate of these regressions are significant at a 5 percent or lower level of significance in a two-tailed t-test. The overall F statistics of the OLS and Log-likelihood ratios for the Tobit models are found to be significant. For the OLS specification, we perform the White test of heteroscedasticity on the residuals obtained from this regression. We reject the null hypothesis of homoscedasticity and correct the standard errors of the estimates using heteroscedasticity consistent standard errors (White, 1980). This correction results in consistent and efficient parameter estimates. We report the heteroscedasticity consistent standard errors in our results.

The next aspect of this paper is to look at subsequent compliance 2, 3, and 4 tax years after the ASFR treatment. The main objective is to estimate the impact of the ASFR treatment, both directly and indirectly, on subsequent voluntary filing compliance. The estimates obtained from the three logistic regressions are expressed in terms of log-odds ratios, which are not very intuitive for interpretation purposes. Therefore, we estimate marginal effects at the sample means for interpretational convenience. The results are reported in Table 5.

TABLE 5. Model 3: Voluntarily Filing a Return Two, Three, and Four Years from TDI Assignment

Dependent Variable: Taxpayer Voluntarily Filed a Tax Return 'j' Tax Years Later; j=2,3 and 4	Two Tax Years After		Three Tax Years After		Four Tax Years After	
	Coefficients	Marginal Effects	Coefficients	Marginal Effects	Coefficients	Marginal Effects
Indicator of ASFR Treatment	0.42 (0.02)***	0.09	0.25 (0.01)***	0.06	0.18 (0.01)***	0.04
Predicted Probability of ASFR Working a Case	0.51 (0.02)***	0.11	0.89 (0.03)***	0.21	1.14 (0.03)***	0.27
Number of Cycles to ASFR Treatment (30-day letter issued)	-0.01 (0.001)***	-0.002	-0.004 (0.0003)***	-0.001	-0.003 (0.0002)***	-0.0006
Self-correction: Taxpayer Filed Return on TDI Prior to Due Date of Tax Return j	1.55 (0.02)***	0.35	1.32 (0.02)***	0.31	1.01 (0.02)***	0.24

Source: Internal Revenue Service Individual Master File Status and Transaction History, and Case Creation Nonfiler Identification Process. Data extracted February 2015.

Notes: Not all explanatory variables shown. See Appendix.

*p<0.1; **p<0.05; ***p<0.01; Marginal Effects are calculated at the sample means.

Based on the estimates reported in Table 5, the direct impact of the ASFR treatment on voluntarily filing a return 2, 3, and 4 tax years after ASFR treatment is positive. The likelihood of filing increases by 9, 6, and 4 percentage points, respectively, keeping other variables fixed. This direct effect appears to be decreasing over time, which seems reasonable. The indirect effects are also positive but larger than the direct effects: 11 percent, 21 percent, and 27 percent, respectively for this period. This is consistent with the hypothesis that the indirect effect is stronger than the direct effect and has lasting effect over the subsequent years. The marginal impact of “cycle to 30 days” declines over time.

The indicator for self-correction is defined as the taxpayer who voluntarily filed his tax return after ASFR treatment but before the next tax return was due. It is a measure that captures the “willingness” of the taxpayer to be compliant post-ASFR treatment. We find a positive and significant impact on future voluntary filing compliance. The marginal impact is found to be strong and ranges from 24 percent to 35 percent during the period of 2 to 4 years. The impact of other control variables on subsequent voluntary filing compliance is reported in the Appendix.

Simulation

It seems intuitive that working more ASFR cases will result in increased tax revenue and more voluntarily filed returns. Both results come from the direct impact of IRS enforcing filing requirements, and indirect effects of the level of enforcement. We use our estimated models of payment and subsequent filing to simulate the impacts of working more of the available nonfiler cases in ASFR.

Using the Tax Year 2009 cases in our study, we estimate the increase in dollars collected and the increase in returns subsequently filed voluntarily in the simulated counterfactual scenario of ASFR working an additional 100,000 cases from the available inventory. To create our counterfactual data (i.e., ASFR working 100,000 more cases), we replicate the Tax Year 2009 data and randomly designate 100,000 of the unworked cases as being treated by ASFR.²⁸ We assume that the initial ASFR letter was issued immediately; thus the number of cycles to 30-day letter would remain as zero. In order to include the indirect effects, P(ASFR) in the subsequent filing model is increased to reflect the increase in the proportion of available inventory worked. Working 100,000

²⁸ We selected a random sample of unworked cases to be conservative in our estimates compared to using the ASFR prioritization criteria to select the next 100,000 unworked cases.

more cases for Tax Year 2009 would increase the proportion of available inventory worked by 0.13. Thus, 0.13 is added to the P(ASFR) for each of the cases in the counterfactual scenario, but it is constrained to be no more than one.²⁹

Let

$E(P_{ai})$ be the predicted dollars collected for taxpayer i based on the actual data;

$E(P_{ci})$ be the predicted dollars collected for taxpayer i based on the counterfactual data;

$P(R_{ai})$ be the predicted probability of filing a return for 2011 based on the actual data; and

$P(R_{ci})$ be the predicted probability of filing a return for 2011 based on the counterfactual data.

For the OLS model of payments the expected payments are calculated as

$$E(P) = X\hat{\beta}.$$

For the Tobit model of payments the expected payments are calculated as

$$E(P) = (X\hat{\beta} + \sigma\lambda)\Phi\left(\frac{X\hat{\beta}}{\sigma}\right).$$

Then, the estimated increase in payments would be

$$\text{Increase in payments} = \sum_{\forall i} (E(P_{ci}) - E(P_{ai})).$$

The increase in returns subsequently filed would be

$$\text{Increase in returns} = \sum_{\forall i} (P(R_{ci}) - P(R_{ai})).$$

The results of the simulated increase in enforcement are reported in Table 6. These estimates are to some degree conservative because we do not estimate with our analysis the indirect effects on the cases that were never in a post notice delinquent return status. It is reasonable to expect that increased enforcement might increase the cases responding to delinquent return notices or may choose to file their return timely. That impact is beyond the scope of this paper. Also, we are assuming that the additional cases are being selected randomly. To a degree, the IRS can prioritize the “best” case to select; therefore, the increases in dollars and/or returns would be somewhat larger.

TABLE 6. Simulated Total Impact of Working 100,000 More ASFR Cases for Tax Year 2009

Model	Total Increase	Increase Per ASFR Case Started
Increase in Payments (Linear Model)	\$118,077,994	\$1,181
Increase in Payments (Tobit Model)	\$326,192,842	\$3,262
Increase in Voluntarily Filed Returns in 2011	19,469	0.19
Increase in Voluntarily Filed Returns in 2012	24,563	0.25
Increase in Voluntarily Filed Returns in 2013	29,166	0.29

Based on data reported in internal CFO cost accounting, the cost per case closed varied between \$53 and \$80 per case for FY 2009 to FY 2013.³⁰ Using the \$80 as an estimated cost per case, revenue collected relative to the cost would be just under 15:1 using the linear model estimate and just over 40:1 using the Tobit model

²⁹ No case had a probability high enough to invoke the constraint.

³⁰ Since ASFR is an automated system, the cost per working an ASFR case remains fairly stable for direct ASFR assignments.

estimate. In addition, every \$110 spent on the ASFR program results in an additional voluntarily filed return ($\$80/ (.19 + .25 + .29)$), based on our model estimates.

Conclusions and Direction for Further Research

In this paper, we develop a model of taxpayer payments of unpaid taxes associated with delinquent returns and the subsequent decision to file future tax returns timely. We focus specifically on the impact of IRS enforcement via the ASFR process. Our model provides estimates of both the direct effects of the IRS enforcement and the indirect effects of that enforcement.

Our estimates suggest significant direct and indirect impacts of enforcing filing compliance via the ASFR process, for both payment and subsequent filing compliance. The indirect effects are somewhat smaller than the direct effects for payment of taxes on delinquent returns. However, the indirect effects on subsequent filing compliance are large relative to the direct effects. In addition, the relative magnitude of the estimates is similar to effects reported in other studies of an audit's effect on reporting compliance. While we do not examine data around cost of an ASFR case, cost estimates from other studies compared to our model estimates suggest that the return on investment is relatively high.

Based on our estimates, it is clear that, given the downward trend in the number of ASFR cases worked, there has been and will be significant declines in enforcement revenue from the ASFR program and decreases in the number of returns voluntarily filed. Clearly, the IRS is devoting fewer resources to the ASFR program. Some of this decline may be the result of shifting resources to other programs that may be more productive or important, and thus would actually result in overall improvement in accomplishing the goals of tax administration. However, it is likely that much of the decline is the result of the IRS response to decreasing real budgets and increasing responsibilities. It may well be that the decreases in the ASFR starts were part of an optimal strategy to absorb the budget shocks. Our estimates suggest the ASFR resource reductions result in both significant decreases in enforcement revenue and reduced voluntary filing compliance.

This paper is a first attempt to estimate the impacts of nonfiling enforcement on the gross and net tax gap. With more years and a broader set of data, this study could be enhanced in many ways. We make some fairly restrictive assumptions about how taxpayers form their expectations of the "likelihood" that IRS will start an ASFR case. Our model assumes taxpayers do not update their expectation of the probability they are selected as new tax years come due, especially when the taxpayer has not yet been selected. Our research could be extended to consider all the nonfiler treatment streams and impact on all taxpayers, including those who have always filed timely or at least have always resolved in the notice process. This would, however, dramatically increase scope and complexity of the analysis.

References

- Allingham, M. G., and Sandmo, A. (1972). "Income tax evasion: A theoretical analysis." *Journal of Public Economics*, 1:3-4, 323-38.
- Alm, J., and Torgler, B. (2011). "Do ethics matter? Tax compliance and morality." *Journal of Business Ethics*, 101(4), 635-651.
- Amemiya, T. (1973). "Regression analysis when the dependent variable is truncated normal." *Econometrica: Journal of the Econometric Society*, 997-1016.
- Andreoni, J., Erard, B., and Feinstein, J. (1998). "Tax compliance." *Journal of Economic Literature*, 818-860.
- Coase, R. H. (1960). "The problem of social cost." *Journal of Law & Economics*, 3, 1.
- Erard, B., and Ho, C. (2001). "Searching for ghosts: Who are the nonfilers and how much tax do they owe?" *Journal of Public Economics*, 81(1), 25-50.
- Greene, W. H. (2008). *Econometric Analysis*. Granite Hill Publishers.
- Hashimzade, N., Myles, G. D., and Tran-Nam, B. (2013). "Applications of behavioural economics to tax evasion." *Journal of Economic Surveys*, 27(5), 941-977.

- Hurst, E., Li, G., and Pugsley, B. (2014). "Are household surveys like tax forms? Evidence from income underreporting of the self-employed." *Review of Economics and Statistics*, 96(1), 19–33.
- Kleven, H. J., Knudsen, M. B., Kreiner, C.T., Pedersen, S., and Saez, E. (2011). "Unwilling or unable to cheat? Evidence from a randomized tax audit experiment in Denmark." *Econometrica*, 79(3), 651–692.
- Lederman, L., and Sichelman, T. M. (2013). "Enforcement as substance in tax compliance." *Washington and Lee Law Review*, 70, 1679–1749.
- Ljungqvist, L., and Sargent, T. J. (2000). *Recursive Macroeconomic Theory*. Granite Hill Publishers.
- Milgrom, P., and Roberts, J. (1992). *Economics, Organization and Management*. Engelwood Cliffs, NJ: Prentice Hall.
- IRS Office of the Chief Financial Officer, Financial Management, Office of Cost Accounting Cost-Based Performance Measures Automated Substitute for Return (ASFR) FY2009—FY2013, Unpublished internal CFO document, 2014.
- Plumley, A. H. (1996). *The determinants of individual income tax compliance: Estimating the impacts of tax policy, enforcement, and IRS responsiveness*. Internal Revenue Service. Publication 1916 (Rev. 11–96).
- White, H. (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica: Journal of the Econometric Society*, 817–838.

Appendix

Complete Model Results

TABLE A1. Probability of ASFR Working a Case, Probit Model

Explanatory Variables*	Parameter Estimate	Standard Error	Wald Statistic	P-value
Intercept	-0.955	0.024	1546.944	<.0001
Log of wages reported on Form W-2	0.006	0.001	55.396	<.0001
Log of Non-Employment Compensation reported on Form 1099-MISC	0.020	0.001	451.147	<.0001
Log of the balance due calculated for a potential Substitute for Return (SFR) return.**	0.051	0.003	336.260	<.0001
Indicator = 1 if the delinquent return is for Tax Year 2007	0.796	0.007	14731.298	<.0001
Indicator = 1 if the delinquent return is for Tax Year 2008	0.750	0.006	13439.036	<.0001

n=277,314

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

*Some explanatory variables have been suppressed.

** To SFR Potential Tax Assessment (IRPSFRTX), add the sum of Advance Earned Income Credit (AEIC), computed tax on premature distributions, and Self-Employment (SE) tax, less the sum of current year credit balance, excess FICA, and withholding. It is set to zero if the result is a negative.

TABLE A2. Predicted Dollars Collected, OLS Regression

Explanatory Variables	Parameter Estimate	Std. Error	t Value	P-value	Heteroscedasticity Consistent Statistics			Tolerance	Variance Inflation
					Std. Error	t Value	Pr > t		
Intercept	-8759.950	156.695	-55.900	<.0001	534.600	-16.39	<.0001		0
Indicator = 1 if ASFR treated the case within 3 years of TDI Assignment	841.834	51.601	16.310	<.0001	48.776	17.26	<.0001	0.39823	2.51114
Probability of ASFR working the case	331.310	94.200	3.520	0.0004	120.122	2.76	0.0058	0.7405	1.35044
For ASFR treated cases, this is the number of weeks from TDI Assignment to ASFR 30-day letter; else = 0	-8.589	0.903	-9.520	<.0001	0.875	-9.82	<.0001	0.43236	2.31287
Indicator = 1 if the taxpayer filed for an extension along with a payment.	1792.805	102.999	17.410	<.0001	302.685	5.92	<.0001	0.96361	1.03776
Indicator = 1 if Federal employee	293.454	70.985	4.130	<.0001	50.888	5.77	<.0001	0.89953	1.11169
Indicator = 1 if the taxpayer had either interest reported on Form 1099-INT or dividends reported on Form 1099-DIV	525.954	41.646	12.630	<.0001	29.672	17.73	<.0001	0.66814	1.49669
Indicator = 1 if the taxpayer had pensions or annuity benefits reported on Form 1099-R	-226.482	40.306	-5.620	<.0001	40.841	-5.55	<.0001	0.7708	1.29736
Indicator = 1 if the taxpayer had unemployment benefits reported on Form 1099-G	-173.513	71.438	-2.430	0.0151	25.913	-6.7	<.0001	0.94645	1.05658
Indicator = 1 if the taxpayer had any withholding reported on information returns received by the IRS.	330.025	55.980	5.900	<.0001	74.708	4.42	<.0001	0.38716	2.58291
Indicator = 1 if the taxpayer reported a Married Filing Joint Filing Status on their prior-year return.	250.959	44.370	5.660	<.0001	62.702	4	<.0001	0.86498	1.1561
Indicator = 1 if the taxpayer had a filing requirement code of not required to file therefore indicating no tax form package be mailed to the taxpayer.	-276.580	43.919	-6.300	<.0001	33.924	-8.15	<.0001	0.83508	1.19749
Log of the balance due calculated for a potential Substitute for Return (SFR) return.*	962.347	18.423	52.240	<.0001	62.981	15.28	<.0001	0.76381	1.30923
Log of wages reported on Form W-2	21.747	5.115	4.250	<.0001	7.347	2.96	0.0031	0.39261	2.54706
Indicator = 1 if the taxpayer had more than two Forms W-2	-536.156	67.825	-7.910	<.0001	54.341	-9.87	<.0001	0.88362	1.13171
Log of the number of information returns received for the taxpayer.	450.948	26.162	17.240	<.0001	35.342	12.76	<.0001	0.54001	1.85181
Log of the total positive income reported on the prior-year return filed.	30.122	3.617	8.330	<.0001	5.444	5.53	<.0001	0.76085	1.31431

n=277,314

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

* To SFR Potential Tax Assessment (IRPSFRTX), add the sum of Advance Earned Income Credit (AEIC), computed tax on premature distributions, and Self-Employment (SE) tax, less the sum of current year credit balance, excess FICA, and withholding. It is set to zero if the result is a negative.

TABLE A3. Predicted Dollars Collected, Tobit Regression Censored at Zero

Explanatory Variable	Parameter Estimate	Standard Error	t Value	P-value
Intercept		556.250	-97.500	<.0001
Indicator = 1 if ASFR treated the case within 3 years of TDI Assignment	11385.000	182.148	62.500	<.0001
Probability of ASFR working the case	8241.278	340.377	24.210	<.0001
For ASFR treated cases, this is the number of weeks from TDI Assignment to ASFR 30-day letter; else = 0	-103.764	3.258	-31.850	<.0001
Indicator = 1 if the taxpayer filed for an extension along with a payment.	5883.410	296.835	19.820	<.0001
Indicator = 1 if Federal employee	1968.786	221.887	8.870	<.0001
Indicator = 1 if the taxpayer had either interest reported on Form 1099-INT or dividends reported on Form 1099-DIV	2935.717	142.671	20.580	<.0001
Indicator = 1 if the taxpayer had pensions or annuity benefits reported on Form 1099-R	1590.049	137.725	11.550	<.0001
Indicator = 1 if the taxpayer had unemployment benefits reported on Form 1099-G	-3073.823	269.743	-11.400	<.0001
Indicator = 1 if the taxpayer had any withholding reported on information returns received by the IRS.	1998.573	195.552	10.220	<.0001
Indicator = 1 if the taxpayer reported a Married Filing Joint Filing Status on their prior-year return.	1365.964	144.111	9.480	<.0001
Indicator = 1 if the taxpayer had a filing requirement code of not required to file therefore indicating no tax form package be mailed to the taxpayer.	-3657.524	175.781	-20.810	<.0001
Log of the balance due calculated for a potential Substitute for Return (SFR) return.*	1902.135	62.939	30.220	<.0001
Log of wages reported on Form W-2	143.076	17.401	8.220	<.0001
Indicator = 1 if the taxpayer had more than two Forms W-2	-4130.596	253.805	-16.270	<.0001
Log of the number of information returns received for the taxpayer.	2306.890	91.902	25.100	<.0001
Log of the total positive income reported on the prior-year return filed.	669.091	12.431	53.830	<.0001

n=277,314

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

* To SFR Potential Tax Assessment (IRPSFRTX), add the sum of Advance Earned Income Credit (AEIC), computed tax on premature distributions, and Self-Employment (SE) tax, less the sum of current year credit balance, excess FICA, and withholding. It is set to zero if the result is a negative.

TABLE A4. Probability of Voluntarily Filing a Subsequent Tax Return Two Tax Years Later, Logistic Regression

Explanatory Variables	Parameter Estimate	Standard Error	Wald Statistic	P-value
Intercept	-1.167	0.043	750.824	<.0001
Probability of ASFR working the case	0.512	0.025	429.598	<.0001
Indicator = 1 if ASFR treated the case prior to the due date of the tax return 2 years later.	0.420	0.019	484.754	<.0001
For ASFR treated cases, this is the number of weeks from TDI Assignment to ASFR 30-day letter; else = 0	-0.009	0.001	230.103	<.0001
Indicator = 1 for non-ASFR treated TDIs if the taxpayer filed their return before the due date of the return due 2 tax years later.	1.554	0.019	6978.168	<.0001
Indicator = 1 if the case is classified as SB/SE	0.144	0.009	256.123	<.0001
Indicator = 1 if the taxpayer filed for an extension on the unfiled return.	0.138	0.011	155.763	<.0001
Indicator = 1 if the taxpayer had interest reported on Form 1099-INT	0.118	0.009	157.036	<.0001
Indicator = 1 if the taxpayer had pensions or annuity benefits reported on Form 1099-R	0.062	0.010	40.752	<.0001
Indicator = 1 if the taxpayer had unemployment benefits reported on Form 1099-G	-0.113	0.018	38.399	<.0001
Indicator = 1 if the taxpayer had any withholding reported on information returns received by the IRS.	0.106	0.014	57.165	<.0001
Indicator = 1 if the taxpayer reported a Married Filing Joint Filing Status on their prior-year return.	0.058	0.011	27.618	<.0001
Indicator = 1 if the taxpayer had wages reported on a Form W-2	0.380	0.014	777.412	<.0001
Indicator = 1 if the taxpayer had a delinquent prior-year tax module that was closed and received a TC59x.	0.860	0.016	3081.011	<.0001
Indicator = 1 if the taxpayer had a zero-dollar tax assessment on their prior-year tax module indicating either no tax due or a Substitute for Return (SFR).	-0.209	0.014	233.139	<.0001
Log of the balance due calculated for a potential SFR return.*	-0.026	0.005	31.367	<.0001
Number of years from the delinquent return and the taxpayer's filed prior-year tax module.	-0.034	0.001	867.929	<.0001
Log of the total positive income reported on the prior-year return filed.	0.027	0.001	583.954	<.0001

n=277,314

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

* To calculate SFR Potential Tax Assessment (IRPSFRTX), add the sum of Advance Earned Income Credit (AEIC), computed tax on premature distributions, and Self-Employment (SE) tax, less the sum of current year credit balance, excess FICA, and withholding. It is set to zero if the result is a negative.

TABLE A5. Probability of Voluntarily Filing a Subsequent Tax Return Three Tax Years Later, Logistic Regression

Explanatory Variables	Parameter Estimate	Standard Error	Wald Statistic	P-value
Intercept	-1.592	0.042	1466.150	<.0001
Probability of ASFR working the case	0.893	0.025	1269.457	<.0001
Indicator = 1 if ASFR treated the case prior to the due date of the tax return 3 years later.	0.251	0.014	308.130	<.0001
For ASFR treated cases, this is the number of weeks from TDI Assignment to ASFR 30-day letter; else = 0	1.323	0.016	6574.608	<.0001
Indicator = 1 for non-ASFR treated TDIs if the taxpayer filed their return before the due date of the return due 3 tax years later.	-0.004	0.000	230.569	<.0001
Indicator = 1 if the case is classified as SB/SE	0.180	0.009	418.407	<.0001
Indicator = 1 if the taxpayer filed for an extension on the unfiled return.	0.186	0.011	297.219	<.0001
Indicator = 1 if the taxpayer had interest reported on Form 1099-INT	0.083	0.009	80.643	<.0001
Indicator = 1 if the taxpayer had pensions or annuity benefits reported on Form 1099-R	0.024	0.010	6.411	0.0113
Indicator = 1 if the taxpayer had unemployment benefits reported on Form 1099-G	-0.140	0.018	62.174	<.0001
Indicator = 1 if the taxpayer had any withholding reported on information returns received by the IRS.	0.115	0.014	70.395	<.0001
Indicator = 1 if the taxpayer reported a Married Filing Joint Filing Status on their prior-year return.	0.069	0.011	40.971	<.0001
Indicator = 1 if the taxpayer had wages reported on a Form W-2	0.465	0.013	1221.670	<.0001
Indicator = 1 if the taxpayer had a delinquent prior-year tax module that was closed and received a TC59x.	0.813	0.015	2816.524	<.0001
Indicator = 1 if the taxpayer had a zero-dollar tax assessment on their prior-year tax module indicating either no tax due or a Substitute for Return (SFR).	-0.212	0.014	246.543	<.0001
Log of the balance due calculated for a potential SFR return.*	0.017	0.004	13.659	0.0002
Number of years from the delinquent return and the taxpayer's filed prior-year tax module.	-0.027	0.001	602.749	<.0001
Log of the total positive income reported on the prior-year return filed.	0.025	0.001	526.521	<.0001

n=277,314

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

* To calculate SFR Potential Tax Assessment (IRPSFRTX), add the sum of Advance Earned Income Credit (AEIC), computed tax on premature distributions, and Self-Employment (SE) tax, less the sum of current year credit balance, excess FICA, and withholding. It is set to zero if the result is a negative.

TABLE A6. Probability of Voluntarily Filing a Subsequent Tax Return Four Tax Years Later, Logistic Regression

Explanatory Variables	Parameter Estimate	Standard Error	Wald Statistic	P-value
Intercept	-1.167	0.043	750.824	<.0001
Probability of ASFR working the case	0.512	0.025	429.598	<.0001
Indicator = 1 if ASFR treated the case prior to the due date of the tax return 4 years later.	0.420	0.019	484.754	<.0001
For ASFR treated cases, this is the number of weeks from TDI Assignment to ASFR 30-day letter; else = 0	-0.009	0.001	230.103	<.0001
Indicator = 1 for non-ASFR treated TDIs if the taxpayer filed their return before the due date of the return due 4 tax years later.	1.554	0.019	6978.168	<.0001
Indicator = 1 if the case is classified as SB/SE	0.144	0.009	256.123	<.0001
Indicator = 1 if the taxpayer filed for an extension on the unfiled return.	0.138	0.011	155.763	<.0001
Indicator = 1 if the taxpayer had interest reported on Form 1099-INT	0.118	0.009	157.036	<.0001
Indicator = 1 if the taxpayer had unemployment benefits reported on Form 1099-G	-0.113	0.018	38.399	<.0001
Indicator = 1 if the taxpayer had any withholding reported on information returns received by the IRS.	0.106	0.014	57.165	<.0001
Indicator = 1 if the taxpayer reported a Married Filing Joint Filing Status on their prior-year return.	0.058	0.011	27.618	<.0001
Indicator = 1 if the taxpayer had wages reported on a Form W-2	0.380	0.014	777.412	<.0001
Indicator = 1 if the taxpayer had a delinquent prior-year tax module that was closed and received a TC59x.	0.860	0.016	3081.011	<.0001
Indicator = 1 if the taxpayer had a zero-dollar tax assessment on their prior-year tax module indicating either no tax due or a Substitute for Return (SFR).	-0.209	0.014	233.139	<.0001
Log of the balance due calculated for a potential SFR return.*	-0.026	0.005	31.367	<.0001
Number of years from the delinquent return and the taxpayer's filed prior-year tax module.	-0.034	0.001	867.929	<.0001
Log of the total positive income reported on the prior-year return filed.	0.027	0.001	583.954	<.0001

n=277,314

Source: IRS, Compliance Data Warehouse Individual Master File Status and Transaction History, and Individual Case Creation as of February 2015 (cycle 201508)

* To calculate SFR Potential Tax Assessment (IRPSFRTX), add the sum of Advance Earned Income Credit (AEIC), computed tax on premature distributions, and Self-Employment (SE) tax, less the sum of current year credit balance, excess FICA, and withholding. It is set to zero if the result is a negative.

3



Improving Tax Administration by Understanding Taxpayer Behavior

Hernandez ♦ Kettle ♦ Ruda ♦ Sanders

Rasey ♦ Wiley

McDaniels ♦ Wooten

Behavioural Interventions in Tax Compliance: Evidence from Guatemala

Stewart Kettle, University of Bristol, *Marco Hernandez*, The World Bank, *Simon Ruda*, The Behavioural Insights Team, London, and *Michael Sanders*, University of Oxford¹

1. Introduction

The collection of tax is a vital function of government, and in most countries tax revenues are the primary means through which public goods and services are funded. The international community has widely recognised that taxation is crucial to ensure sustainable development and to allow independence from foreign aid in the long term (Mascagni (2015)). Guatemala collects one of the lowest shares of public revenues, with government revenues amounting to about 12 percent of gross domestic product (GDP) during the period 2011-2014. This is well below the average of 26 percent in Latin America and the world average of 32 percent (World Bank (2014a)).

The country's low tax revenue to GDP ratio is partially attributable to tax evasion. For the income tax, the estimated average rate of tax evasion for individuals and firms in Guatemala is about 64 percent, compared to an average of about 50 percent for Latin America (Gómez Sabaini and Jimenez (2012)). Low public revenues limit the capacity of this middle-income State to deliver services and eradicate poverty. More than half of Guatemala's 15 million people live below the poverty line, the coverage of basic goods and services is modest, and human development indicators are weak (World Bank (2014b)). Reducing tax evasion is consequently a high priority for the government of Guatemala.

We designed and implemented a randomised controlled trial (RCT) using reminders to promote tax compliance. Individuals and businesses that had failed to declare their income tax for the 2013 Tax Year were randomly assigned to one of six treatment arms. The motivation behind this experiment was to see if tax reminders sent by the Guatemalan Tax Authority (*Superintendencia de Administración Tributaria*, SAT) could be effective at increasing tax declarations and payment in Guatemala. Individuals were assigned to receive the original letter from SAT or one of four letters adapted using behavioural design, or to a control group. The four adapted letters were simplified, made the desired action clear, and included the address of the declaration website. Three of the letters included additional persuasive messages: one included a social norms message, another highlighted nondeclaration as a deliberate choice, and the third emphasised national pride.

The results show that whilst all letters were successful at increasing the rate of declaration relative to the control group, only two of the letters were successful at increasing the rate of payment and the average amount paid conditional on payment. The original letter composed by SAT helped to increase declaration rates but did not have a statistically significant effect on the rate of tax payment or the amount paid relative to the control group.

The two best performing letters were the social norms and deliberate choice letters. These letters increased the rate of payment as well as the average amount paid conditional on payment, overall more than tripling tax receipts. The social norms and deliberate choice letters increased the average amount paid per taxpayer by \$13.97 (210 percent) and \$17.95 (269 percent), respectively, relative to no letter.² These point estimates are also significantly different from the original letter and the behavioural letter, the latter of which was identical to these conditions other than the two sentences added for the social norms and deliberate choice messages. These sentences alone increased the average amount paid by \$6.69 and \$10.78 over and above the impact of the basic behavioural letter (however these are not significantly different from each other).

These results provide evidence that, in this context, increasing the moral cost by referring to a social norms condition has a significant impact on payment. This impact is achieved with a social norm of only 64.5 percent, supporting

¹ Stewart Kettle and Michael Sanders are also affiliated with the Behavioural Insights Team. The views expressed in this paper are those of the authors and should not be attributed to the institutions they represent.

² Figures in United States Dollars are calculated based on an exchange rate of 0.131 dollars per Guatemalan Quetzal.

our hypothesis that due to the perception of chronic tax evasion in Guatemala this relatively small percentage still causes taxpayers to update their beliefs about the rate of compliance upwards. The results also show that framing the decision to evade as a deliberate choice can have a significant impact over and above the reminder and deterrent message. This message is aimed at removing inaction as a strategy to avoid blame and thereby remove the tendency to remain with the status quo of inaction (Hallsworth (2013)). Local average treatment effects (LATE) show that the deliberate choice letter, when received, increases the average amount paid by \$23.05 compared with no letter, highlighting that the wording of the letter has a very substantial impact on behaviour.

We estimate that the best performing letter, the deliberate choice letter, if sent to all taxpayers in the sample, would have generated an estimated \$757,837 of extra tax revenue in 11 weeks compared to no letter, over thirty-five times the cost of sending the letters, in addition to reducing administrative costs for the tax authority by encouraging taxpayers to pay earlier. Conversely, the original letter would have cost the tax authority \$21,090 to send to the whole sample, but we find no evidence that this letter increased the average amount paid.

Our sample includes individual and business taxpayers and so, unlike previous studies in the literature, we can estimate the impact of the same behavioural messaging on both types of taxpayers. The results are similar for individuals and businesses. This is the first such result, to our knowledge, that shows a positive impact on tax payment of increasing the moral cost for businesses.

This paper makes several contributions to the existing literature on tax compliance. First, it presents a national level randomised evaluation exploring the effects of tax reminders on declaration, rate of payment, and payment amount. Second, to our knowledge the study represents the first randomised evaluation of tax reminders that investigates the same treatments on both individuals and firms. Third, it provides evidence on the effectiveness of a social norm message where the norm is only marginally a norm at 64.5 percent. Fourth, in industrialised countries, eventual compliance with tax regulations is nearly 100 percent, and so interventions, such as those described in Hallsworth *et al.* (2014), primarily have the effect of bringing forward tax compliance. The results of this trial remain significant after 12 months and suggest behavioural messages may be more advantageous in developing countries with low levels of tax compliance because they could help generate additional tax revenues, rather than just increase the speed of payment. Fifth, we show that the intervention also causes habituation; taxpayers who receive the two best performing letters are more likely to pay the same tax the following year with no further reminder. Overall, the study provides a compelling case in favour of the application of behavioural sciences to public policy, and for rigorous testing of small adaptations to communications written by policymakers or government administrators.

2. Background

The past decade has seen an increase in applied research on the behaviour and decision making of taxpayers (Slemrod and Weber (2012); Hallsworth (2014)). The majority of this research has been conducted in industrialised countries, with the impact of messages tested through the medium of tax letters. RCTs have been conducted in Australia, Argentina, Austria, Chile, Denmark, Germany, Israel, Peru, Switzerland, the United States, the United Kingdom, and Venezuela, among others (Blumenthal *et al.* (2001); Torgler (2007); Wenzel and Taylor (2004); Kleven *et al.* (2011); Ariel (2012); Castro and Scartascini (2013); Del Carpio (2013); Fellner *et al.* (2013); Ortega and Sanguinetti (2013); Pomeranz (2013); Dwenger *et al.* (2014); and Hallsworth *et al.* (2014)). The literature provides evidence that reminders, deterrent messages, and moral and social norms messages can affect tax behaviour, at least in some contexts.

In an overview of RCTs in this field to date, Hallsworth (2014) finds that most interventions that focus on increasing the perceived probability of enforcement action, or the perceived severity of this action, increase subsequent compliance. The magnitude of effect sizes are, however, frequently small. Of the 22 studies in the review that include non-deterrence messages, only half showed a significant effect on compliance.

Four recent applied studies from Latin America demonstrate the range of findings in this literature. First, in a letter trial where 10,228 individuals were sent information about property tax payment in three municipalities in Lima, Peru, both a simple reminder message and a social norms message were found to be effective at increasing payment. However no impact was found for an enforcement message net of the reminder effect (Del Carpio (2013)). Second, in a trial involving 23,000 participants and property tax payment in a municipality in Argentina, Castro and Scartascini (2013) found that *only* a deterrence message was effective and that social norms and public messaging were ineffective. Third, Ortega and Sanguinetti (2013) find no statistically significant effects of any messages sent to businesses regarding a business tax in a municipality in Venezuela, net of a reminder effect. However, due to a relatively small sample (N=6,000) and a large number of tests (5 conditions), the trial did not have the statistical power to detect effect sizes of

magnitudes similar to those found elsewhere. Fourth, in a study involving over 400,000 firms in Chile that looked at the role of third party information, Pomeranz (2013) found that deterrent letters threatening audit not only increased value-added tax (VAT) payments for the firms themselves, but also did so for their clients.

The range of findings of these interventions is likely due to two main factors: heterogeneity in context (recipients, tax type, trust in institutions, social norms and tax culture), and heterogeneity in the intervention messaging. Variations in the specific presentation of information, beyond the particular element that is meant to be tested, could explain the difference in the findings in the literature. For instance, Del Carpio (2013) presents information about compliance levels with a bar graph, but uses a pie chart to present information about enforcement levels. As Hasseldine (2000) notes, the effectiveness of a treatment may depend on seemingly insignificant elements of its presentation, highlighting the importance of isolating interventions and rigorous testing of small adaptations to how messages are framed in particular contexts.

In the behavioural literature in general, and tax compliance more specifically, there is limited evidence on the impact of behavioural messages on businesses. We note that studies involving businesses have tended to focus on enforcement messaging (see for example Ortega and Sanguinetti (2013), and Pomeranz (2013)). Nonetheless, a trial conducted in Israel targeting 4,395 corporations, in which the intervention focused on the societal costs of noncompliance, found no effect on sales values or payments, and actually documented an increase in VAT deductions (Ariel (2012)). Our paper is to our knowledge the first study with the same treatments presented to individuals and businesses.

3. A Simple Theoretical Framework

In this section we construct a framework to structure our thinking on tax compliance behaviour and the mechanisms through which the policy tools we will later test may work. First we develop a model that combines two theoretical approaches to understanding tax compliance: the deterrence and nondeterrence approaches. Individuals in this model maximise their utility subject to the expected monetary costs, as well as moral costs, of compliance or noncompliance. Based on these assumptions, the framework shows that the policy tools available to governments are to increase the perceived likelihood of punishment, the perceived cost of being punished if caught, and the moral costs of noncompliance or the moral benefit of compliance. Finally, we relax the assumption of full attention to these costs and consider that individuals may not be fully conscious of the trade-off that they are facing, as their tax return is not salient, suggesting the policy tool of reminder messages (nudges) of how and when to pay tax.

The original economic model used to analyse tax compliance was developed by Allingham and Sandmo (1972). The model sees taxpayers as rational utility maximisers and suggests that a taxpayer's decision to pay or evade tax is based on the trade-off between the monetary cost of complying and the expected cost of evading. A common objection to this model is that it cannot explain high levels of compliance given the low probability of evasion being detected. Proponents of the model generally respond that the taxpayer's decision is based on the perceived risk of noncompliance rather than the actual risk, or that the classical model needs to recognize the endogeneity of enforcement, the pervasiveness of matching tax returns against third-party information, and the deterrence value of the additional time and monetary costs necessitated when one is subject to enforcement.

A competing perspective on tax compliance is that citizens are fundamentally predisposed to cooperate. This “nondeterrence” approach contends that taxpayers' decisions are not based on monetary payoff maximisation alone, but are also influenced by moral and social norms (Andreoni *et al.* (1998); Kirchler (2007); Erard and Feinstein (1994); and Torgler (2007)). Proponents of the nondeterrence theory criticise the deterrence model for being overly individualistic, and neglectful of the reality that taxpayers are embedded in a wider web of interactions and identities (Alm (2012); and Taylor (2003)). From this perspective, traditional deterrence activities may create an adversarial relationship in which taxpayers react against the attempt to control their actions, which in turn undermines voluntary compliance (Brehm and Brehm (1981); Hessing *et al.* (1992); and Sheffrin and Triest (1992)). Nevertheless, proponents of the nondeterrence approach rarely go as far as arguing that deterrence approaches should be neglected completely.

A more nuanced position is that both deterrence and nondeterrence approaches can affect tax behaviour (Smith (1992); Kirchler (2007); and Hallsworth (2014)), and we take this argument as a starting point for our framework. We follow the approach of Gordon (1989) that nondeterrence factors can be integrated into the utility function of the individual as “moral costs,” preserving the theoretically tractable assumption of utility maximisation while, we contend, increasing the real-world applicability. For simplicity, the following framework assumes the taxpayer is risk-neutral and that fines are fixed. We also assume for simplicity linear preferences for income, but our conclusions are robust to alternative functional forms. We present the utility functions of paying (u_p) and evading tax (u_e) respectively as follows:

$$u_p = y - t + \sigma\alpha$$

$$u_e = y - pf - \mu\delta,$$

where y is the income of the individual, t is his tax liability (here we assume it is fixed for simplicity), α is the utility produced from paying the tax, σ is the extent to which the moral benefit is transparent, bounded between 0 and 1 (see Gabaix and Laibson (2005a)), p is the probability of getting a fixed fine of f , δ is the moral disutility produced from not paying the tax, and μ is the extent to which the moral cost is transparent. In the framework, the decision to pay is based on the taxpayer's income minus her fixed tax liability plus the moral utility that the taxpayer gets from paying her tax. The decision to evade is the probability of getting fined multiplied by her income minus the fine, plus the probability of not getting fined multiplied by her income, minus the moral disutility of evading.

Based on the framework above, the taxpayer will pay her tax if the following condition holds:

$$pf + \sigma\alpha = \mu\delta > t$$

That is, a taxpayer will pay her tax if the probability of getting fined multiplied by the fine, plus the moral benefit she gets from paying, plus the moral cost she derives from not paying, is greater than her tax liability. Accordingly, the framework predicts that one way to deter tax evasion is through increased monitoring of taxpayers, and increased sanctions for noncompliance—the policy tools traditionally favoured by tax authorities (Feld and Larsen (2012)). The effect on net government revenue is ambiguous, however, as increasing the perception of enforcement activity may be costly. An alternative for government is to increase the moral benefit from compliance, increase the moral cost of noncompliance, or increase the salience of moral costs, hence making failing to comply relatively less attractive for the taxpayer. These aims could be achieved by making the immorality of not paying more salient or by reminding taxpayers of their own moral identity—for example by increasing information on social norms.

Finally, we relax the assumption of full attention to these costs and consider that individuals may not be conscious of the trade-off they are facing, as their tax return is not salient. Given the infrequency with which tax returns need to be filed, and the number of other tasks in people's lives, it is possible that participants who are late in filing their return, may not even be considering the trade-off they are facing, as their tax return is simply not salient. This could be a case of forgetting or procrastination. In this context, reminders can provide information to taxpayers that they have not paid and serve as a nudge towards payment (Thaler and Sunstein (2008)). If this is the case, even a generically worded letter (or any other form of reminder) might be effective at increasing declarations by forcing participants to consider the trade-off of not complying. We therefore add a reminder effect to the list of policy tools that may increase compliance.

The framework presented is not comprehensive, but it seeks to illustrate the theoretical concepts underlying our treatments. As such, we have chosen to present only the simplest case of a binary decision in this model, although it could be expanded to cover more complex cases without loss of generality. In summary, we have considered five mechanisms for increasing tax compliance: (1) increasing the perceived probability of getting caught; (2) increasing the perceived punishment if caught; (3) increasing the moral costs of noncompliance, such as by making the decision to evade a deliberate choice; (4) increasing the moral benefit from compliance by providing information on public goods and social norms; and (5) providing reminders (nudges) on how, where, and when to pay.

4. Experimental Context and Design

4.1 Guatemala

Guatemala is the largest economy in Central America with a GDP of about \$53 billion and a per-capita gross national income of about \$4,880 in 2013, putting it the top half of the 'lower middle income' country classification (World Bank (2014b)). Human development indicators, however, are weak compared to other middle-income countries, and this is associated with modest, and often unequal, coverage of basic goods and services. As a percentage of total government spending, Guatemala's health and education expenditures are below those of other Latin American countries, which is mainly due to the government's limited revenue intake. An analysis of the determinants of growth in Guatemala highlights that reducing poverty and social inequality would require a more comprehensive social program for which additional fiscal resources are needed (World Bank (2014a)).

Partly due to weak tax collection, the Government of Guatemala's public revenue (and spending) as a percentage of GDP are among the lowest in the world, which limits its capacity to fund social programs for the poor. Though Guatemalan officials have prioritized tax reform, approving tax legislation has historically been difficult (see World

Bank (2014a)). During the period 2000–2014, government revenues in Guatemala averaged about 12 percent of GDP (of which about 90 percent were tax revenues), placing Guatemala in 186th place out of 187 countries, according to the 2014 International Monetary Fund's World Economic Outlook database.³ Unlike other countries with low levels of tax collection, like Mexico or Panama, Guatemala does not possess significant nontax revenues. The size of the informal or shadow economy in Guatemala is estimated to be around 52 percent of GDP during the period 1999–2006—well above the average of 36 percent estimated for 98 developing countries by Schneider *et al.* (2010).

4.2 Tax Regime Selection

In Guatemala there are two regimes for income tax: a profits tax, which taxpayers can voluntarily opt into at the beginning of the tax year, and a gross income tax, which is the default option and is associated with simplified accounting standards. For the 2013 Tax Year, the profits tax entails a tax rate of 31 percent (28 percent in 2014) on taxable profits, and the gross income tax regime entails a 6-percent tax on gross revenue.

The trial focuses on taxpayers who have self-selected into the profits tax regime. Tax in this regime is calculated on the basis of the accounting profit adjusted for nondeductible expenses, exempt income, and foreign-source income. The sample is therefore likely to consist of a subset of taxpayers with either high gross income or with substantial deductibles (such as income from capital). The reasons for choosing to study this tax regime choice were twofold. First, the profits tax is subject to an annual inventory and tax return (subject to quarterly declarations and payments), whereas for the gross income tax regime tax is declared and paid monthly. This made the prospect of increasing compliance among taxpayers under this regime particularly attractive for SAT, as any effects identified in this trial would be more economically significant. Second, the profits regime involves deductions and exemptions, and therefore more evasion opportunities than the income tax regime based on gross revenue. The tax was due to be declared, and paid, by the end of March and end of April 2014, respectively.

In Guatemala there are also currently two tax regimes for the VAT, both subject to a 12-percent tax on the value of products. These regimes differ only in eligibility (based on total sales) and how and when the tax is paid. In Table 1 we present information on all tax declarations made in 2014 for the four tax regimes described for comparison.

TABLE 1. Number of Declarations and Amount Declared by Tax Type, Fiscal Year 2014

Tax Regime	Number of Declarations	Percentage of Declarations	Mean Tax Liability (\$)	Percent of Tax Receipts
VAT Small Taxpayers	4,044,887	49.56%	7.12	1.80%
VAT General	2,540,366	31.13%	374.66	59.67%
Gross Income	1,290,015	15.81%	259.41	20.98%
Profits Income	286,029	3.50%	978.06	17.54%

Source: Authors' calculations based on data received from the Guatemalan Tax Authority.

Table 1 shows that the profits income and gross income regimes produce similar total tax receipts. The number of declarations is much lower for profits income because, as described above, the gross income regime is declared monthly, whereas profits income is declared quarterly. Similarly, the VAT small taxpayers declare monthly whereas the VAT general is declared quarterly. Overall, the income tax produces less income for the State than the VAT.

4.3 Sample Selection and Random Treatment Assignment

The trial involved taxpayers who were due to declare their annual income tax for Fiscal Year 2013, but had failed to do so. The sample was taken from the universe of taxpayers due to pay their annual income tax under the “Regime over Profits from Lucrative Activities” for 2013 according to SAT (n=115,999). From this universe, the sample includes all taxpayers who had failed to declare by May 16, 2014 (n=44,952). Finally a number of taxpayers (n=1,565) were removed from the sample as they had already been contacted by SAT about failing to pay this tax, leading to a final sample of n=43,387 taxpayers.⁴ The 1,563 taxpayers already contacted (and hence removed from the sample) were chosen by the

³ Since 2000, Guatemala's government revenues have fallen from 12.3 percent of GDP to 11 percent in 2014, while in Bangladesh, which on average ranked 187th during the entire period, government revenues increased from 8.3 percent of GDP to 10.5 percent during the same period.

⁴ Two taxpayers were removed from the sample during the analysis stage. These taxpayers were both extreme outliers and noncompliers (i.e., did not receive the letters). The participants paid an amount over 500 times more than the average amount paid, and therefore caused convergence problems in a number of specifications.

tax authority based on their calculations as to those most likely to respond to letters and with the highest tax revenue potential. Assuming that the tax authority is accurate in its targeting, this creates a selection issue as the taxpayers in our sample are systematically harder to reach than the general population, and/or have the lowest tax revenue potential. This implies that our results may provide a conservative estimate of the impact of the trial letters on the universe of taxpayers that initially failed to declare.

Randomisation was conducted at the individual level with no stratification. Following Bruhn and McKenzie (2009) we chose to use a pure random draw rather than testing for balance and re-randomising in the event of failure of balance on observables. Following Kahan *et al.* (2014) any significant differences in observables between treatment arms are controlled for in the analysis.

4.4 Experimental Design and Treatments

Taxpayers were randomly assigned to one of six arms: a control arm who did not receive a tax reminder, and five treatment arms. Sample sizes per arm are reported in Table 2. More participants were allocated to the control group than to any of the treatment arms due to logistical restrictions on the number of letters that could be sent out by the Guatemalan mail office. The treatments included the original SAT letter and four adapted versions of the letter. Table 2 presents a summary of the letter types, and subsequently we describe the letters in more detail. The original Spanish versions of the letters are included in the Appendix.

TABLE 2. Summary of Letter Types

Group	Description
Control Group <i>n</i> = 12,397	<ul style="list-style-type: none"> ■ No Letter
Original SAT Letter <i>n</i> = 6,198	<ul style="list-style-type: none"> ■ Simple reminder to declare, no information on how to declare
Behavioural Design Letter <i>n</i> = 6,197	<ul style="list-style-type: none"> ■ Begins with a short 'call to action' stating declaration is needed now ■ Specified the website address to declare at ■ Informed taxpayers that they can pay in instalments ■ Included the deterrent message: <i>"If you do not declare you may be audited and face the procedure established by law"</i>
Behavioural Design + Social Norm Letter <i>n</i> = 6,198	<ul style="list-style-type: none"> ■ The same as the behavioural design letter but including the message: <i>"According to our records, 64.5 percent of Guatemalans declared their income tax for the year 2013 on time. You are part of the minority of Guatemalans who are yet to declare for this tax"</i>
Behavioural Design + Deliberate Choice Letter <i>n</i> = 6,198	<ul style="list-style-type: none"> ■ The same as the behavioural design letter but including the message: <i>"Previously we have considered your failure to declare an oversight. However, if you don't declare now we will consider it an active choice and you may therefore be audited and could face the procedure established by law"</i>
Behavioural Design + National Pride Letter <i>n</i> = 6,199	<ul style="list-style-type: none"> ■ The same as the behavioural design letter but excluding the call to action and the deterrent message. ■ Softer tone, including an image of the Guatemalan flag and the phrase: <i>"You are a Guatemalan citizen and Guatemala needs you. Be a good citizen and submit the 2013 annual return of income tax ... Are you going to support your country?"</i>

4.4.1 Original SAT Letter (*n* = 6,198)

This group was sent the standard letter used by the Guatemalan Tax Authority. A range of different letter formats is sent to taxpayers; for this arm, SAT used the format most commonly used to remind taxpayers of their income tax debt. The letter has a number of positive attributes. For example, it is succinct enough to fit onto one page. There is, however, a tendency for official letters to be written from the perspective of government officials or civil servants, and the SAT letter is no exception to this. For example, the SAT letter includes the legal basis for the communication, and although it includes information saying that the taxpayer needed to declare, it does not include information on *how* to declare. We therefore argue that this letter serves as a basic nudge and that the impact of this letter will show us the propensity to declare and to pay with a simple reminder.

4.4.2 Behavioural Design Letter (n = 6,197)

Making clear the key required actions is an essential requirement for any letter that requests the recipient to complete an action. Thus, the original SAT letter was adapted to focus on concisely explaining what the recipient has done (not declare her tax), what she needs to do now (declare her income tax), and how to do it (via *Declaraguante*, SAT's online tax declaration portal).⁵ Specifically, the original SAT letter was adapted using behavioural design, based on three principles: (1) making the messages and required actions simple and clear; (2) introducing persuasive messages; and (3) checking tone and overall 'feel'. This "Behavioural Design Letter" served as a basis for other treatments, as further specified below.

Furthermore, the adapted letter aimed to communicate the main message of the letter as quickly as possible (see Gabaix and Laibson (2005b)), using an eye catching "call to action" at the start of the letter which states:

"Please file your declaration of income tax"

The purpose of this "call to action" is to make the desired action more salient, so that taxpayers can understand the key action required of them at only a glance of the letter. Nonessential text to this purpose was eliminated with the goal of keeping the letters as simple as possible. To make the desired action more likely, all of the behaviourally designed letters include the website address of the Tax Authority, *Declaraguante*. The web address was included as a bullet point in order to catch the eye of the reader.

The only additional text included in the letter was information that taxpayers can pay off their income tax debt in instalments once they have declared online. Although close to 100 percent of payment for this tax type can be done online, paying in instalments is possible only in person at SAT offices, and less than 1 percent of taxpayers use this option. It is likely that this is due to a lack of information on this service available to taxpayers, so this message was included to rectify this.

In this letter an additional deterrent message was included as a reason to declare:

"If you do not declare, you may be audited and could face the procedure established by law."

The aim of this message is to increase the perception of the taxpayers that they will be audited or face legal proceedings if they fail to comply, thus increasing the 'cost' of noncompliance.

4.4.3 Behavioural Design + Social Norm (n = 6,198)

This letter is the same as the behavioural design letter with the addition of a persuasive message that describes a social norm in the first paragraph of text:

"According to our records, 64.5% of Guatemalans declared their income tax for the year 2013 on time. You are part of the minority of Guatemalans who are yet to declare for this tax"

This message aims to increase the "moral benefit" of compliance by emphasising the social norm of compliance under this tax regime. We note that the social norm in this case is only 64.5 percent, and consequently too low to be classed a norm. Previous research on social proof offers evidence that norms as low as 70-75 percent can be effective in motivating behaviour; for example Gerber and Rogers (2009) showed that stating that 71 percent of citizens voted in a previous election increased voter turnout, and Goldstein *et al.* (2008) found that a norm stating that 75 percent of hotel users reused their towels increased towel reuse. Due to the fact that it would be unethical to cite false social norms, there is limited evidence currently on the cutoff level for when social norms are effective. However, in a recent experiment, Anik *et al.* (2014) test the level of contingent matches for online charitable donations that are the most effective. In the experiment, individuals are offered to have their recurring donations matched if "X" percent of donors agree to on that day. Thus, the authors can manipulate the level to find out the most effective level of social proof. The authors find that the most effective level is 75 percent, however this is only against 25 percent, 50 percent and 100 percent, rather than subtler values such as 65 percent. In our experiment, we test the value of 64.5 percent, a much lower social norm. We hypothesise that due to the perceived culture of noncompliance in Guatemala, a norm of 64.5 percent will still update individuals' perception of the social norm upwards and thereby increase the moral cost of noncompliance.

Following Wenzel (2005) the social norm message is worded to make the norm as specific to the recipient taxpayers as possible by referring only to Guatemalans who declare under the same tax regime. Further, we follow Hallsworth

⁵ During the course of the trial, close to 100 percent of taxpayers in Guatemala used *Declaraguante* to declare the income tax.

et al. (2014) by explicitly stating that the message recipient is not complying with the norm and is thus in the minority. These framings are all aimed at reducing the level of abstraction or “psychological distance” to the norm, as elaborated in construal level theory (Trope and Liberman, 2010; Hallsworth *et al.* 2014).

4.4.4 Behavioural Design + Deliberate Choice (n = 6,198)

This letter is again the same as the behavioural design letter; however this version includes a message framing dishonesty as a deliberate choice:

“Previously we have considered your failure to declare an oversight. However, if you don’t declare now we will consider it an active choice and you may therefore be audited and could face the procedure established by law.”

Most real decisions have a status quo alternative, that is, doing nothing or maintaining one’s current or previous decision. A large literature of experiments show that individuals tend to disproportionately stick with the status quo (Samuelson and Zeckhauser (1988); and Anderson (2003)). This is due to two reasons: first, losses from acting may loom larger than gains (Tversky and Kahneman (1981)); and second, moral violations tend to be judged less wrong when the violation results from inaction as opposed to action (DeScioli *et al.* (2012)). Individuals may use this greater ambiguity and uncertainty associated with acts of omission in order to minimise future psychic costs arising from the threat to self-image of acting dishonestly (Hallsworth (2013)).

This message aims to overcome status quo bias (i.e., doing nothing) by framing the dishonest behaviour as a deliberate choice. By notifying the reader that her failure to comply will be treated as a deliberate choice, this message aims to eliminate omission as an excuse for noncompliance now, thus increasing cognitive dissonance of taxpayers’ beliefs and increasing the perceived cost of payment in the future. The wording also gives the taxpayer an exemption for not previously declaring, which introduces an element of reciprocity, as the taxpayer is given the sense that he has been granted a favour. The text is also worded to give the impression that the behaviour of the taxpayer is being closely monitored and serves to increase the perception of punishment for noncompliance.

4.4.5 Behavioural Design + National Pride (n = 6,199)

This letter does not include the deterrent message or the “call to action.” Instead, the whole tone of the letter is ‘softer’ and is meant to invoke the national responsibility of paying taxes. The letter includes an image of the Guatemalan flag and includes a message to encourage identification as a taxpaying citizen:

“You are a Guatemalan citizen and Guatemala needs you. Be a good citizen and submit the 2013 annual return of Income Tax. [...] Are you going to support your country?”

Similar to the social norms letter this message aims to increase the ‘moral benefit’ of compliance, however it does so by emphasising the national benefit of tax payment. The letter includes no deterrent message so as not to crowd out intrinsic motivation for tax declaration (Brehm and Brehm (1981); Hessing *et al.* (1992); and Sheffrin and Triest (1992)).

4.5 Outcome Variables

We have four outcome measures of interest, each of which were measured both 11 weeks and 52 weeks after the letters were sent on May 29, 2014:

- 1) Declaration (binary): defined as the taxpayer completing online declaration (taxpayers declare their income for the year to the tax authorities through the *Declaraguante* website);
- 2) Payment (binary): defined as the taxpayer paying a nonzero amount of tax;
- 3) Payment (log) amount (continuous): the total amount paid by the taxpayer conditional on payment; and
- 4) Payment amount: the total amount paid by the taxpayer unconditional on payment.

The outcome variables were all collected regularly and automatically as administrative data by SAT. The declaration data were collected automatically from the *Declaraguante* website, and the payment data were sent to SAT by the relevant banks involved in processing tax payments.

Our data contain the anonymised tax records of the 43,387 participants in this trial. For each taxpayer, we are able to identify his region (North, South, West and Central), whether the taxpayer is identified as an individual or a

business, and our outcomes of interest, measured at the time intervals specified. The term *individual* includes a relatively heterogeneous group, including professionals and small business owners (for example traders). Businesses are defined as registered ‘legal entities’ and are required by law to have an accountant. Additionally, we have data from the mail delivery service, which records whether or not a letter was delivered for each taxpayer.

Table 3 shows descriptive statistics for the experiment, both for the full sample and broken down by treatment condition. Also reported in the table are the results of simple balance checks. The majority of participants are in the Central region (as are the majority of Guatemalans), and a minority (25.9 percent) are businesses. Baseline characteristics do not vary significantly across our treatment groups.

TABLE 3. Descriptive Statistics and Balance Checks

Condition	Number of Observations	Number of Letters Delivered	Percent Central Region	Percent Businesses	Baseline ¹ Mean Payment (\$)
Full Sample	43,387	21,781	67.4	25.9	61.5
No Letter (control group)	12,397	0	67.0	25.6	60.3
Original SAT Letter	6,198	4,382	67.6	25.8	66.0
Behavioural Letter	6,197	4,316	67.9	25.7	60.4
Behavioural + Social Norms	6,198	4,332	67.2	25.9	55.6
Behavioural + Deliberate Choice	6,198	4,337	67.5	26.3	68.6
Behavioural + National Pride	6,199	4,414	67.2	26.0	59.3

¹ Mean payment in 2012 (prior to trial).

5. Results

5.1 Estimation

Our primary estimation consists of an Intention To Treat (ITT) analysis of the impact of the letters. The main regressions use data collected on August 14, 2014, some 11 weeks after the letters were sent out. We also use data collected after 12 months to test for longer-term effects.

We estimate the impact of our treatments on four outcome variables:

1. **declaration**—a binary variable set to 1 if taxpayer i declares her tax by the above dates and 0 otherwise;
2. **payment**—a binary variable set to 1 if taxpayer i pays a positive amount of tax and 0 otherwise;
3. the (log) **amount paid** by the taxpayer—that is, the amount paid **conditional on payment**; and
4. the **amount paid unconditional on payment**.

The first three models are evaluated by ordinarily least squares (OLS) multiple regression, and the fourth is evaluated by Poisson Pseudo Maximum Likelihood (PPML) following Santos Silva and Tenreyto (2006). The PPML model has been shown to outperform OLS and Tobit models with many zero observations in the data and in the presence of heteroskedasticity. Tobit would lead to biased estimates unless the standard errors are homoskedastic and normal. For each of these outcome variables we estimate the parameters of the following model:

$$Y_i = z + \beta_1 T_{1i} + \beta_2 T_{2i} + \beta_3 T_{3i} + \beta_4 T_{4i} + \beta_5 T_{5i} + \gamma \varphi_i + U_i$$

Where: Y_i is the outcome variable; z is a constant; T are binary variables representing the five treatment conditions outlined above (T_1 = Original SAT letter, T_2 = Behavioural Insights, T_3 = Behavioural Insights + Social Norm, T_4 = Behavioural Insights + Deliberate Choice, T_5 = Behavioural Insights + National Pride); φ is a vector of control variables comprising taxpayers’ characteristics; β and γ are parameters to be estimated; and U_i is the error term. In addition to testing for differences between the control group and the treatment letters we also test for differences between the different letter arms. In the OLS regressions we control for heteroscedasticity in U_i using robust standard errors.

5.2 ITT Parameter Estimates

Table 4 shows the parameter estimates for the ITT model described above. Overall, the best performing conditions are the social norms and deliberate choice letters. These letters are shown to increase the rate of declaration, rate of payment, as well as the amount paid, both conditional and unconditional on payment. On the other hand, whilst the original letter composed by SAT is successful at increasing the rate of declaration it has no impact on the rate of payment or the average amount paid.

Turning first to the rate of declaration, the point estimates show that the original letter composed by SAT increases declaration rates by 3.6 percentage points, a 31 percent increase relative to the control group. The most successful letter for increasing declaration rates—the deliberate choice letter—increases declarations by 5.4 percentage points (a 46 percent increase) over the control group, and 1.8 percentage points over the original SAT letter (a 10 percent increase). We note that of the behavioural letters, only the deliberate choice performs significantly better than the original SAT letter ($p = 0.005$).

TABLE 4. ITT Estimates of Treatment Impacts on Tax Declaration and Payment, at 11 Weeks

	(1) Declared	(2) Paid	(3) LogAmount	(4) Amount ^λ
Original Letter	0.036 *** (0.006)	0.004 (0.003)	-0.042 (0.122)	0.148 (5.406)
Behavioural Letter	0.043 *** (0.006)	0.005 (0.003)	0.362 ** (0.122)	7.527 * (3.299)
Behavioural + Social Norms	0.048 *** (0.006)	0.017 *** (0.003)	0.309 ** (0.113)	13.972 *** (3.946)
Behavioural + Deliberate Choice	0.054 *** (0.006)	0.014 *** (0.003)	0.307 ** (0.115)	17.953 ** (6.235)
Behavioural + National Pride	0.038 *** (0.006)	0.010 ** (0.003)	0.173 (0.117)	9.491 * (4.677)
Business	0.008 (0.004)	0.019 *** (0.002)	0.255 ** (0.078)	10.617 ** (3.709)
Large Taxpayer	0.517 *** (0.098)	0.462 *** (0.058)	1.086 (0.573)	43.093 *** (7.333)
Paid in 2012	0.000 *** (0.000)	0.000 *** (0.000)	0.000 *** (0.000)	0.002 *** (0.000)
North	-0.033 *** (0.006)	-0.012 *** (0.003)	-0.114 (0.142)	-8.785 (5.355)
South	0.010 (0.006)	0.015 *** (0.003)	0.150 (0.110)	4.498 (4.583)
West	-0.019 *** (0.006)	-0.005 (0.003)	0.058 (0.128)	3.508 (5.706)
Constant	0.118 *** (0.004)	0.032 *** (0.002)	2.658 *** (0.085)	
Observations	43,387	43,387	2,008	43,387

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported (dy/dx).

For payment, the differences between the letters are more pronounced. The results show that the original SAT letter does not have a statistically significant impact on the rate of payment, or the average amount paid. That is, the original letter helped to increase the rate at which taxpayers filed their declarations but did not affect payment. This is consistent with an interpretation of our hypothesis that the original SAT letter provides a mere information effect for participants who have forgotten the need to file their return, but that people who are most likely to forget are those who owe no money, for whom there is a much smaller moral cost to not declaring. The adapted letters fare better with respect to tax payments. The three behavioural letters with persuasive messages increase the rate of payment significantly compared to the control condition, and all four of the behavioural letters, with the exception of the national pride condition are shown to increase the average amount paid conditional on payment.

Of the behavioural prompts, the best performing letters are the social norms condition and the deliberate choice condition. These two letters significantly increase the rate of payment (by 1.7 and 1.4 percentage points, respectively) and the average amount paid conditional on payment (by 43.6 and 38.5 percent, respectively). These impacts on the rate of payment and the amount paid combined add up to a large impact on the average amount paid per letter sent. The estimates in column 4 show that each letter sent increased the average unconditional amount paid by \$13.97 (210 percent) and \$17.95 (269 percent), up from a control group average of \$6.70. While these effects are not significantly different from each other, they are significantly different from the other conditions. These magnitudes are large and provide compelling evidence that the behavioural adaptations to the letters have had a large impact on payment behaviour. Both of these conditions triple the impact of the basic behavioural letter on the average amount paid by taxpayers. We note that in the case of the social norms letter this is achieved with a social norm of only 64.5 percent, supporting our hypothesis that due to the perception of chronic tax evasion in Guatemala this still updates taxpayers' beliefs of compliance upwards. Similarly, the results show that framing a failure to declare as a deliberate choice helped to promote tax compliance.

5.3 *Treatment on the Treated*

Letters were hand-delivered by a mailing service. In some cases, letters were not delivered. Reasons for this included the address not existing or the taxpayer having moved. Since treatment was not complied with in all cases, we are likely to have a downward biased estimate of the impact of letters on people who received them. On the other hand, complier average causal effects (CACE) are biased upwards, as those assigned to treatment but not reached by the mail office have a significantly lower propensity to declare and pay than those in the control group.⁶ Next we use treatment conditions as instrumental variables to estimate the impact of receiving a letter on our outcome variables, that is, the local average treatment effects (LATE).

Table 5 shows the results from our two-stage least squared (2SLS) estimations. The results show that all of the treatment effects that were significant in the ITT analysis remain significant, and no other treatments have become significant. We find that, as expected, the magnitude of the impact of the letters has increased, and the order of magnitude of the point estimates remains unchanged. For declaration, again only the deliberate choice condition is significantly different from the original letter ($p = 0.004$). The magnitudes of the impacts on declaration are large: for people who actually received the deliberate choice letter it increased declaration by 7.7 percentage points over no letter, a 68-percent increase. The impact of the deliberate choice letter on declaration over the original SAT letter is 2.6 percentage points, a 17-percent increase.

The three behavioural letters with additional messages are effective at increasing the rate of payment. The results show that the impact of the most successful letter, the social norms letter, increases payment by 2.4 percentage points; a 67-percent increase in payment. Again, this is significantly different from the original letter ($p = 0.001$) and the behavioural letter ($p = 0.002$) but not significantly different from the deliberate choice letter. This provides further evidence that the social norms message and the deliberate choice message are effective at increasing the rate of payment, whereas the original and simple behavioural letters are not. Even if it is assumed that the lack of significance of the impact of the behavioural letter is a power issue, the two sentence changes in the social norms and deliberate choice conditions show that both letters have triple the impact of the other letters on the rate of payment.

The LATE estimates for the impact of our treatments on the amount paid show that the social norms letter and deliberate choice letter increase the amount paid on average by \$18 ($p < 0.001$) and \$23 ($p < 0.001$) respectively, compared to \$6.67 in the control condition—that is, an increase of 273 percent and 346 percent, respectively. The LATE estimates provide further evidence that the original SAT letter has no impact on the rate of payment or the average amount paid.

5.4 *Cost-Benefit Analysis*

We perform a cost-benefit analysis of our treatments using as a basis the results from our ITT estimation on the amount declared, since this is likely to be the most relevant to policymakers. The total cost of sending letters to all 43,387 taxpayers in our sample is estimated to be \$21,090. Given that the original SAT letter was found to have no statistically significant impact on amount declared, this letter therefore represents a cost of \$21,090 with no increased revenue for

⁶ Complier Average Causal Effects (CACE) and Defier Average Causal Effects (DACE) are not presented but are available on request.

the tax authority. Nonetheless, this letter was found to increase declaration, which may bring other benefits because it gives the tax authority more information about taxpayers.

TABLE 5. 2SLS Estimates of Treatment Impacts on Tax Declaration and Payment, at 11 Weeks

	(1) Declared	(2) Paid	(3) LogAmount	(4) Amount ^λ
Original Letter	0.051 *** (0.008)	0.006 (0.005)	-0.048 (0.149)	-1.655 (15.310)
Behavioural Letter	0.062 *** (0.008)	0.007 (0.005)	0.442 ** (0.149)	10.899 * (5.191)
Behavioural + Social Norms	0.069 *** (0.008)	0.024 *** (0.005)	0.383 ** (0.140)	18.218 ** (5.790)
Behavioural + Deliberate Choice	0.077 *** (0.008)	0.020 *** (0.005)	0.389 ** (0.145)	23.047 ** (8.457)
Behavioural + National Pride	0.053 *** (0.008)	0.015 ** (0.005)	0.207 (0.141)	13.106 * (6.482)
Business	0.013 *** (0.004)	0.021 *** (0.002)	0.275 *** (0.079)	12.274 ** (4.050)
Large Taxpayer	0.547 *** (0.098)	0.470 *** (0.058)	1.231 * (0.575)	49.919 *** (13.062)
Paid in 2012	0.000 *** (0.000)	0.000 *** (0.000)	0.000 *** (0.000)	0.002 *** (0.000)
North	-0.022 *** (0.006)	-0.010 ** (0.003)	-0.083 (0.142)	-6.729 (5.105)
South	0.022 *** (0.006)	0.018 *** (0.003)	0.184 (0.110)	7.418 (4.702)
West	-0.018 ** (0.006)	-0.005 (0.003)	0.055 (0.128)	4.469 (5.575)
Constant	0.113 *** (0.004)	0.031 *** (0.002)	2.643 *** (0.088)	
Observations	43,387	43,387	2,008	43,387

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported.

Concerning benefits, the average amount paid in the control group was \$6.67. The best performing letter was the deliberate choice letter, which increased the amount declared by \$17.95 for each letter sent. If this letter were sent to all 43,387 taxpayers in the sample, we estimate that this would have generated \$778,927—meaning a profit of \$757,837. This represents a 36-times return on investment for the tax authority. As this letter was not sent out to all taxpayers, we also estimate the revenue generated by the trial itself. The trial itself brought in an estimated \$303,366 at a cost of \$15,065 thereby generating \$288,301 for SAT.

5.5 Businesses and Individuals

The empirical focus so far has been on average treatment effects. We now turn to analysis of heterogeneous treatment effects by estimating whether individuals respond differently to the interventions than do businesses (registered legal entities, as defined above). Governments and nongovernment organizations invest considerable resources in trying to alter the behaviour of both individuals and businesses. Despite this, the overwhelming majority of previous studies in behavioural science are concerned almost solely with the behaviour of individuals. Whereas previous studies have considered the behaviour of businesses, no studies of which we are aware compare the effect of the same interventions on both businesses and individuals, with the same outcome measure. Tax compliance, and this trial in particular, offers an opportunity to investigate this.

Table 6 shows the impact of the letters on individuals. The results are broadly similar to the full sample results, the main difference being that the deliberate choice condition is not shown to significantly affect amount declared ($p=0.099$), although we note that the direction is positive and the trial was not designed with power for this subgroup analysis.

TABLE 6. ITT Estimates of Treatment Impacts on Individuals, at 11 Weeks

	(1) Declared	(2) Paid	(3) LogAmount	(4) Amount ^λ
Original Letter	0.038 *** (0.006)	0.006 (0.004)	0.082 (0.138)	6.685 (3.922)
Behavioural Letter	0.043 *** (0.006)	0.006 (0.004)	0.369 ** (0.138)	5.959 * (2.809)
Behavioural + Social Norms	0.047 *** (0.006)	0.015 *** (0.004)	0.311 * (0.130)	11.049 ** (3.891)
Behavioural + Deliberate Choice	0.052 *** (0.006)	0.016 *** (0.004)	0.284 * (0.130)	12.169 (7.384)
Behavioural + National Pride	0.033 *** (0.006)	0.008 * (0.004)	0.329 * (0.137)	9.751 * (4.559)
Large Taxpayer	0.088 (0.358)	0.196 (0.200)	2.298 (1.569)	24.094 *** (6.154)
Paid in 2012	0.000 *** (0.000)	0.000 *** (0.000)	0.000 *** (0.000)	0.001 *** (0.000)
North	-0.037 *** (0.006)	-0.012 *** (0.003)	-0.201 (0.143)	-9.962 (5.439)
South	0.004 (0.006)	0.013 *** (0.003)	0.141 (0.111)	-0.654 (2.841)
West	-0.026 *** (0.006)	-0.007 * (0.003)	0.027 (0.131)	-0.870 (5.065)
Constant	0.120 *** (0.004)	0.032 *** (0.002)	2.622 *** (0.090)	
Observations	32,150	32,150	1,299	32,150

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported.

Table 7 shows the impact of the letters on businesses. The average amount paid by businesses in the control group is \$10.85. The social norms and deliberate choice conditions increase this by \$27.60 and \$37.10, respectively. It is therefore overall even more worthwhile for the tax authority to send behavioural reminders to businesses rather than individuals.

TABLE 7. ITT Estimates of Treatment Impacts on Businesses, at 11 Weeks

	(1) Declared	(2) Paid	(3) LogAmount	(4) Amount ^λ
Original Letter	0.030 ** (0.011)	-0.002 (0.007)	-0.337 (0.240)	-18.103 (14.504)
Behavioural Letter	0.042 *** (0.011)	0.001 (0.007)	0.330 (0.238)	15.713 (9.918)
Behavioural + Social Norms	0.051 *** (0.011)	0.022 ** (0.007)	0.303 (0.213)	27.596 ^λ (10.825)
Behavioural + Deliberate Choice	0.061 *** (0.011)	0.009 (0.007)	0.380 (0.223)	37.097** (14.120)
Behavioural + National Pride	0.051 *** (0.011)	0.018 ^λ (0.007)	-0.072 (0.217)	14.307 (13.833)
Large Taxpayer	0.551 *** (0.105)	0.481 *** (0.070)	0.812 (0.695)	87.775*** (20.520)
Paid in 2012	0.000 *** (0.000)	0.000 *** (0.000)	0.000 *** (0.000)	0.003 *** (0.001)
North	-0.013 (0.017)	-0.017 (0.011)	0.366 (0.398)	-0.470 (12.409)
South	0.044 ^λ (0.018)	0.025 ^λ (0.012)	0.134 (0.307)	24.128 (15.458)
West	0.025 (0.017)	0.005 (0.011)	0.197 (0.333)	18.957 (16.012)
Constant	0.119 *** (0.007)	0.052 *** (0.004)	2.991 *** (0.144)	
Observations	11,237	11,237	709	11,237

Standard errors in parentheses; ^λ $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported.

5.6 Long-Term Effects

In most industrialised countries, eventual compliance with tax regulations is nearly 100 percent, and so interventions, such as those described in Hallsworth *et al.* (2014), primarily have the effect of speeding up tax compliance. In non-industrialised countries with higher rates of tax evasion, behavioural messages may be more advantageous if they generate additional tax revenues. In order to test whether our treatments were primarily having an effect on bringing forward tax compliance or getting individuals to declare what otherwise wouldn't have been declared, we collect further data on our outcome variables after 12 months. This analysis is an extension to the original trial protocol, so we present these results separately.

Table 8 shows the results of our ITT model after 12 months, comparable to the results after 11 weeks in Table 4. The results show that after 12 months the impact of the letters has increased in absolute terms for almost all of our treatments on declaration, rate of payment, and amount declared, other than a few exceptions. One exception is that the behavioural letter without additional messages is now not shown to have a positive impact on the rate of payment or the amount declared. The point estimate for the impact of the deliberate choice conditions on the amount declared has also increased marginally to \$19.74. We estimate that after 12 months the deliberate choice letter would bring in an estimated \$835,370 of revenue for the tax authority if it were sent to the whole sample.

TABLE 8. ITT Estimates of Treatment Impacts at Twelve Months

	(1) Declared	(2) Paid	(3) LogAmount	(4) Amount ^λ
Original Letter	0.037*** (0.006)	0.005 (0.004)	0.002 (0.071)	4.494 (4.782)
Behavioural Letter	0.037*** (0.006)	0.002 (0.004)	0.106 (0.072)	5.894 (3.460)
Behavioural + Social Norms	0.046*** (0.006)	0.018*** (0.004)	0.109 (0.068)	13.602** (4.208)
Behavioural + Deliberate Choice	0.049*** (0.006)	0.012** (0.004)	0.132 (0.069)	19.741** (6.482)
Behavioural + National Pride	0.035*** (0.006)	0.012* (0.004)	0.032 (0.069)	10.196* (4.674)
Business	0.033*** (0.004)	0.046*** (0.003)	0.158*** (0.047)	15.568*** (4.037)
Large Taxpayer	0.184 (0.107)	0.411*** (0.080)	1.006* (0.457)	54.302*** (9.485)
Paid in 2012	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.002*** (0.000)
North	-0.045*** (0.006)	-0.023*** (0.005)	-0.112 (0.083)	-12.290* (5.442)
South	0.002 (0.006)	0.010* (0.005)	0.157* (0.071)	7.801 (4.861)
West	-0.032*** (0.006)	-0.020*** (0.005)	0.140 (0.080)	6.664 (5.638)
Constant	0.153*** (0.004)	0.075*** (0.003)	2.902*** (0.049)	
Observations	43,387	43,387	4,056	43,387

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported.

Table 9 presents results from a LATE analysis after 12 months. The results show similar point estimates to our results after 11 weeks in Table 5. Overall, it can be seen that the main results of the trial remain highly significant after 12 months. These results provide evidence that behavioural messages may be even more advantageous in developing countries with high levels of tax evasion because they may help generate additional tax revenues, rather than just bring payments forward.

TABLE 9. 2SLS Estimates of Treatment Impacts on Tax Declaration and Payment at Twelve Months

	(1) Declared	(2) Paid	(3) LogAmount	(4) Amount ^λ
Original Letter	0.052*** (0.008)	0.007 (0.006)	0.003 (0.088)	6.822 (7.870)
Behavioural Letter	0.053*** (0.009)	0.003 (0.006)	0.133 (0.090)	9.014 (5.006)
Behavioural + Social Norms	0.066*** (0.009)	0.026*** (0.006)	0.138 (0.086)	18.289** (5.599)
Behavioural + Deliberate Choice	0.071*** (0.009)	0.017** (0.006)	0.167 (0.087)	25.545** (8.241)
Behavioural + National Pride	0.050*** (0.008)	0.016 [†] (0.006)	0.038 (0.084)	14.154 [†] (6.150)
Business	0.038*** (0.004)	0.047*** (0.003)	0.165*** (0.047)	17.374*** (4.311)
Large Taxpayer	0.211 [†] (0.107)	0.418*** (0.080)	1.062 [†] (0.458)	62.191*** (13.830)
Paid in 2012	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.002*** (0.000)
North	-0.035*** (0.006)	-0.021*** (0.005)	-0.100 (0.083)	-10.006 (5.219)
South	0.013 [†] (0.006)	0.012** (0.005)	0.172 [†] (0.071)	10.954 [†] (4.919)
West	-0.030*** (0.006)	-0.020*** (0.005)	0.143 (0.080)	7.555 (5.568)
Constant	0.150*** (0.004)	0.074*** (0.003)	2.897*** (0.051)	
Observations	43,387	43,387	4,056	43,387

Standard errors in parentheses; [†] $p < 0.05$, ^{**} $p < 0.01$, ^{***} $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported.

Table 10 presents results at 12 months with alternative outcome measures. The table shows the impact of our treatments on the following year's tax declaration and payment, with no further reminder. That is, the results examine whether the letters cause habituation in tax payments. The results show that the three behavioural messages with additional messages all increase the likelihood of declaration the following year, and our two best performing letters increase the likelihood of payment the following year. The point estimates show that the two best performing letters increase the rate of payment the following year by 0.7 percentage points (a 16-percent increase) with no further reminder. This provides evidence of the letters causing habitual payment for some taxpayers, thus further increasing the overall impact of the letters. None of the letters are found to increase the average amount paid per letter the following year. However, this is to be expected since the trial is not powered to detect a relatively small increase in the amount paid (such as of a proportional magnitude to the impact of the treatments on the binary rate of payment variable).

TABLE 10. LATE Estimates of Treatment Impacts on Declaration the Following Year at Twelve Months

	(1) Declared 2014	(2) Paid 2014	(3) LogAmount 2014	(4) Amount ^λ 2014
Original Letter	0.012 (0.006)	0.000 (0.003)	-0.138 (0.179)	21.300 (17.839)
Behavioural Letter	0.011 (0.006)	-0.000 (0.003)	0.163 (0.180)	-12.559 (12.485)
Behavioural + Social Norms	0.016 [*] (0.006)	0.007 [*] (0.003)	0.043 (0.171)	38.917 (26.714)
Behavioural + Deliberate Choice	0.017 ^{**} (0.006)	0.007 [*] (0.003)	0.194 (0.170)	50.390 (28.959)
Behavioural + National Pride	0.024 ^{***} (0.006)	0.002 (0.003)	0.132 (0.177)	4.209 (12.931)
Business	0.080 ^{***} (0.005)	0.015 ^{***} (0.002)	1.493 ^{***} (0.122)	95.614 ^{**} (29.132)
Large Taxpayer	0.347 ^{**} (0.113)	0.125 [*] (0.058)	1.279 (1.415)	92.943 ^{**} (34.633)
Paid in 2012	0.000 ^{***} (0.000)	0.000 ^{***} (0.000)	0.000 ^{***} (0.000)	0.004 ^{***} (0.001)
North	-0.026 ^{***} (0.006)	0.001 (0.003)	0.084 (0.187)	-28.658 (18.137)
South	0.033 ^{***} (0.007)	0.019 ^{***} (0.003)	-0.228 (0.164)	-27.325 (14.830)
West	0.002 (0.006)	0.009 ^{**} (0.003)	0.477 ^{**} (0.174)	-12.823 (13.258)
Constant	0.176 ^{***} (0.004)	0.035 ^{***} (0.002)	2.835 ^{***} (0.124)	
Observations	43,387	43,387	1,989	43,387

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

^λ Amount (\$) estimated by PPML, marginal effects reported.

6. Conclusion

We have presented the results from a nationwide randomised controlled trial on tax compliance using tax reminders to promote tax compliance in Guatemala, a low-middle-income country with one of the lowest tax revenues in the world relative to the size of the economy, and with a high rate of tax evasion. Taxpayers (businesses and individuals) were randomly allocated to receive either no reminder letter, the letter originally used by the Guatemalan Tax Authority, or four letter variants adapted using behavioural design.

The best performing treatments were a deterrent message framing nondeclaration as an intentional and deliberate choice, rather than oversight (designed to overcome status quo bias), and a social norms message, which referred to the 64.5 percent of taxpayers who had already paid this tax (designed to nudge taxpayers to join the status quo). These letters increased the rate of payment by 1.7 and 1.5 percentage points, respectively, compared to the control condition in which 3.9 percent of participants paid. These letters also increased the average amount paid conditional on paying. These effects combined mean that each of these letters increases the average unconditional amount paid by \$13.97 (210 percent) and \$17.95 (269 percent), per taxpayer respectively, from a control group average of \$6.70. These effects are not statistically significantly different from each other, but are significantly different from the other conditions, including the almost identical behavioural letter.

The trial presented in this paper is to our knowledge the first national level randomised evaluation exploring the effects of tax reminders on declarations, rate of payment, and payment amount. A key contribution of this study is to show that social norms and the deliberate choice message can be effective at increasing both declaration and payment in the context of a developing country, for both individuals and firms. The results after 12 months are also suggestive that the letters are effective at increasing tax revenue rather than just bringing it forward, as is often the case in industrialised countries. The cost-benefit analysis of the trial shows that the social norms and deliberate choice letters give

a 35 times return on investment, and we estimate that if the deliberate choice letter had been sent to the whole sample it would have generated over \$750,000 in additional revenue after 11 weeks, and close to \$850,000 after 12 months. The results also show that the two best-performing letters increase the likelihood that taxpayers will both declare and pay the following year with no further reminder. Overall, the results provide a compelling case for the application of behavioural sciences to public policy, and for rigorous testing of small adaptations to communications written by policymakers or administrators.

References

- Allingham, M. G., and Sandmo, A. (1972). "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 1, 323–328.
- Alm, J. (2012). "Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Experiments, and Field Studies." *International Tax and Public Finance*, 19, 54–77.
- Anderson, C. J. (2003). "The psychology of doing nothing: forms of decision avoidance result from reason and emotion." *Psychological Bulletin*, 129(1), 139.
- Andreoni, J.; Erard, B.; and Feinstein, J. (1998). "Tax Compliance." *Journal of Economic Literature*, 36, 818–860.
- Anik, L., Norton, M., and Ariely, D. (2014). "Contingent Match Incentives Increase Donations." Working Paper.
- Ariel, B. (2012). "Deterrence and Moral Persuasion Effects on Corporate Tax Compliance: Findings From a Randomized Controlled Trial." *Criminology*, 50, 27–69.
- Blumenthal, M.; Christian, C.; and Slemrod, J. (2001). "Do Normative Appeals Affect Tax Compliance? Evidence From a Controlled Experiment In Minnesota." *National Tax Journal*, 54, 125–136.
- Brehm, S. S., and Brehm, J. W. (1981). *Psychological Reactance: A Theory of Freedom and Control*, Academic Press, New York.
- Bruhn, M., and McKenzie, D. (2009). "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1(4), 200–232.
- Castro, L., and Scartascini, C. (2013). "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment." *Inter-American Development Bank Working Paper Series*. Washington, D.C.: Inter-American Development Bank.
- Coleman, S. (1996). "The Minnesota Income Tax Compliance Experiment: State Tax Results." Mimeo.
- Del Carpio, L. (2013). "Are the Neighbours Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru." Princeton University Working Paper.
- DeScioli, P.; Asao, K.; and Kurzban, R. (2012). "Omissions and byproducts across moral domains." *PLOS ONE*, 7(10), e46963.
- Dolan, P.; Hallsworth, M.; Halpern, D.; King, D.; Metcalfe, R.; and Vlaev, I. (2012). "Influencing Behaviour: The Mindspace Way." *Journal of Economic Psychology*, 33, 264–277.
- Duflo, E.; Glennerster, R.; and Kremer, M. (2008). "Using Randomization in Development Economics Research: A Toolkit." In: Schultz, T. Paul, and Strauss, John (Eds.) *Handbook of Development Economics*, Vol. 4, pp. 3895–3962.
- Dwenger, N.; Kleven, H.; Rasul, I.; and Rincke, J. (2014). "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence From A Field Experiment In Germany." Mimeo.
- Economist* (2013). "Guatemala: Edging Back From the Brink." January 26, 2013. Print Edition.
- Economist* (2014). "Tax in Latin America: Burden Sharing." January 20, 2014. <http://www.economist.com/blogs/americasview/2014/01/tax-latin-america?zid=293&ah=e50f636873b42369614615ba3c16df4a>.
- Erard, B., and Feinstein, J. S. (1994). "The Role of Moral Sentiments and Audit Perceptions in Tax Compliance." *Public Finance-Finances Publiques*, 49, 70–89.
- Feld, L. P., and Larsen, C. (2012). "Self-Perceptions, Government Policies and Tax Compliance in Germany." *International Tax and Public Finance*, 19, 78–103.
- Fellner, G.; Sausgruber, R.; and Traxler, C. (2013). "Testing Enforcement Strategies in the Field: Threat, Moral Appeal, and Social Information." *Journal of the European Economic Association*, Vol. II, Issue 3, pp. 634–660.
- Gabaix, X., and Laibson, D. (2005a). "Shrouded attributes, consumer myopia, and information suppression in competitive markets." (No. w11755). National Bureau of Economic Research.

- Gabaix, X., and Laibson, D. (2005b). "Bounded Rationality and Directed Cognition." Unpublished Paper.
- Gerber, A. S., and Rogers, T. (2009). "Descriptive social norms and motivation to vote: Everybody's voting and so should you," *The Journal of Politics*, 71 (January), 178–191.
- Goldstein, Noah J.; Cialdini, R. B.; and Griskevicius, V. (2008). "A room with a viewpoint: Using social norms to motivate environmental conservation in hotels," *Journal of Consumer Research*, 35 (October), 472–482.
- Gómez Sabaini, J.C., and Jimenez, J.P. (2012). "Tax Structure and Tax Evasion in Latin America," *Macroeconomía del Desarrollo series*, No. 118 (LC/L.3455-P). Santiago, Chile: Economic Commission for Latin America and the Caribbean (ECLAC), February.
- Gordon, J. P. F. (1989). "Individual Morality and Reputation Costs as Deterrents to Tax Evasion." *European Economic Review*, 33, 797–805.
- Hallsworth, M. (2013). "The Omission Bias." Unpublished Paper.
- Hallsworth, M. (2014). "The Impact of Field Experiments to Address Tax Compliance." Unpublished Paper.
- Hallsworth, M.; List, J. A.; Metcalfe, R. D.; and Vlaev, I. (2014). "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance" (No. W20007). National Bureau of Economic Research.
- Hasseldine, J. (2000). "Using Persuasive Communications To Increase Tax Compliance: What Experimental Research Has (and Has Not) Told Us." *Australian Tax Forum*, 15, 227–242.
- Hessing, D.; Elffers, H.; Robben, H.; and Webley, P. (1992). "Does Deterrence Deter? Measuring the Effect of Deterrence on Tax Compliance in Field Studies and Experimental Studies." In: Slemrod, J. (Ed.) *Who Pays Their Taxes And Why? Tax Compliance and Enforcement*. Ann Arbor, Michigan: University of Michigan Press.
- Kahan, B. C.; Jairath, V.; Doré, C. J.; and Morris, T. P. (2014). "The Risks and Rewards of Covariate Adjustment In Randomized Trials: An Assessment of 12 Outcomes From 8 Studies." *Trials*, 15(1), 1–7.
- Kahneman, D. (2011). *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kirchler, E. (2007). *The Economic Psychology of Tax Behaviour*, Cambridge University Press.
- Kirchler, E.; Hoelzl, E.; and Wahl, I. (2008). "Enforced Versus Voluntary Tax Compliance: The 'Slippery Slope' Framework." *Journal of Economic Psychology*, 29, 210–225.
- Kleven, H. J.; Knudsen, M. B.; Kreiner, C. T.; Pedersen, S.; and Saez, E. (2011). "Unwilling or Unable To Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica*, 79, 651–692.
- Mascagni, G. (2015). "Tax Experiments in Developing Countries: A Critical Review and Reflections on Feasibility." Centre for Development Impact, *CDI Practice Paper No. 11*, March 11, 2015.
- McGraw, K. M., and Scholz, J. T. (1991). "Appeals to Civic Virtue Versus Attention to Self-Interest—Effects on Tax Compliance." *Law and Society Review*, 25, 471–498.
- OECD (2014), *Revenue Statistics in Latin America (2014)*, OECD Publishing.
- Ortega, D., and Sanguinetti, P. (2013). "Deterrence and Reciprocity Effects on Tax Compliance: Experimental Evidence from Venezuela." Working Paper.
- Pomeranz, D. (2013). "No taxation without information: Deterrence and self-enforcement in the value added tax." (No. w19199). National Bureau of Economic Research.
- Samuelson, W., and Zeckhauser, R. (1988). "Status quo bias in decision making." *Journal of Risk and Uncertainty*, 1(1), 7–59.
- Santos Silva, J. M. C., and Tenreyro, S. (2006). "The Log of Gravity." *The Review of Economics and Statistics*, 88(4), pp. 641–658.
- Schneider, F., Buehn, A., and Montenegro, C. (2010). "Shadow Economies All Over the World: New Estimates for 162 Countries from 1999 to 2007," Policy Research Working Paper 5356. Washington, DC: The World Bank, July.
- Sheffrin, S. and Triest, R. (1992). "Can Brute Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance." In: Slemrod, J. (Ed.) *Why People Pay Taxes: Tax Compliance And Enforcement*. Ann Arbor: University of Michigan Press.
- Slemrod, J., and Weber, C. (2012). "Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy." *International Tax and Public Finance*, 19(1), 25–53.
- Smith, K. W. (1992). "Reciprocity and Fairness: Positive Incentives for Tax Compliance." In: Slemrod, J. (Ed.) *Why People Pay Taxes: Tax Compliance and Enforcement*, Ann Arbor: University of Michigan Press.

- Taylor, N. (2003). "Understanding Taxpayer Attitudes Through Understanding Taxpayer Identities." In: Braithwaite, V. (Ed.) *Taxing Democracy: Understanding Tax Avoidance and Tax Evasion*. Aldershot, UK: Ashgate.
- Thaler, R. H., and Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven: Yale University Press.
- Torgler, B. (2007). *Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis*, Cheltenham, UK: Edward Elgar Publishing.
- Trope, Y., and Liberman, N. (2010). "Construal-Level Theory of Psychological Distance." *Psychology Review*, 117(2): 440–463.
- Tversky, A., and Kahneman, D. (1981). "The Framing of Decisions and the Psychology of Choice." *Science*, 211(4481), 453–458.
- Wenzel, M. (2005). "Misperceptions of Social Norms about Tax Compliance: From Theory to Intervention." *Journal of Economic Psychology*, 26, 862–883.
- Wenzel, M., and Taylor, N. (2004). "An experimental evaluation of tax-reporting schedules: A case of evidence-based tax administration." *Journal of Public Economics*, 88, 2785–99.
- World Bank (2014a). "Guatemala Economic DNA: Harnessing Growth." Washington, DC: The World Bank.
- World Bank (2014b). "World Development Indicators." Washington, DC: The World Bank.

Acknowledgments

This project would not have materialised without the collaboration and commitment shown by the Superintendencia de Administración Tributaria, Guatemala's Tax Authority. We thank the institution and the team led by Mario Figueroa and Hugo Roldán for their support, patience, time, and resources. We also thank Erlend Berg, Anne Brockmeyer,

Michael Hallsworth, and Patricia Lucas for their helpful comments on the paper, and Tania Diaz for her assistance with the randomisation. We acknowledge funding from the World Bank for project implementation.



ALERTA SAT

Guatemala, 23 de mayo de 2014

Chimaltenango, Chimaltenango

Por este medio me permito informarle que de conformidad con los registros con que cuenta para el efecto la Administración Tributaria, usted omitió la presentación de la declaración siguiente:

PERÍODO	IMPUESTO	FORMULARIO
2013	ISR ANUAL	SAT-1411

Por lo antes indicado, se le requiere revisar sus registros y presentar la declaración omitida, haciendo uso de los medios que la SAT ha puesto a disposición de los contribuyentes.

Lo anterior no limita el ejercicio de la potestad con que cuenta la Administración Tributaria para iniciar el procedimiento de determinación correspondiente.

En caso de duda, cordialmente le invitamos a acudir a la oficina tributaria de su conveniencia, donde será atendido y orientado por un analista especializado.

Base legal: artículo 32 de la Constitución Política de la República; artículo 3 del Decreto 1-98 Ley Orgánica de la SAT; artículos 89, 98, 112 inciso c), 112 numeral 6, del decreto 6-91 del Congreso de la República, Código Tributario.

Atentamente,

Dr. Erick René Reyes Lombardo
Superintendente de Administración Tributaria
Superintendencia de Administración Tributaria

SUPERINTENDENCIA DE ADMINISTRACIÓN TRIBUTARIA

Appendix



ALERTA SAT-1

Acatenango, Chimaltenango

Guatemala, 26 de mayo de 2014

Por favor presente su declaración del impuesto sobre la renta

Estimado contribuyente:

Hemos revisado nuestros archivos y encontramos que usted no ha presentado su declaración anual del Impuesto sobre la Renta correspondiente al año 2013.

Si usted no declara, puede ser auditado y ser sujeto al procedimiento establecido por ley.

Por favor declare en un plazo de 10 días después de recibida esta alerta, lo cual será verificado.

- Para ello, visite: <http://declaraguat.gt> (Formulario SAT-1411)

Si usted no puede pagar el monto total ahora, podría ser posible que pague en cuotas, luego de preparar su declaración y acercarse a una oficina de la SAT.

Si tiene alguna consulta, contáctenos a través del teléfono 2329-7111. En ese caso haga referencia a su número de alerta: ALERTA SAT-1

Estaremos verificando cómo responde a esta carta.

Atentamente,

Lic. César Alfredo Leroz Estrada
Gerente Regional Central
Superintendencia de Administración Tributaria
SAT

Base Legal: Artículos 98,112 y 146 del Código Tributario;
Artículo 3 de la Ley Orgánica de la Superintendencia de Administración Tributaria.

Examples of the Control Letter and Letter Variants of the Experiment

**SAT**
SUPERINTENDENCIA DE ADMINISTRACIÓN TRIBUTARIA

ALERTA SAT

Chimaltenango, Chimaltenango

Guatemala, 26 de mayo de 2014

Por favor presente su declaración del impuesto sobre la renta

Estimado contribuyente:

Según nuestros registros, 64.5% de los guatemaltecos declararon a tiempo su Impuesto sobre la Renta correspondiente al año 2013. Usted es parte de una minoría de guatemaltecos que no ha presentado su declaración de este impuesto.

Si usted no declara, puede ser auditado y ser sujeto al procedimiento establecido por ley.

Por favor declare en un plazo de 10 días después de recibida esta alerta, lo cual será verificado.

- Para ello, visite: <http://declaraguatemala.gt> (Formulario SAT-1411)

Si usted no puede pagar el monto total ahora, podría ser posible que pague en cuotas, luego de preparar su declaración y acercarse a una oficina de la SAT.

Si tiene alguna consulta, contáctenos a través del teléfono 2329-7111. En ese caso haga referencia a su número de alerta: ALERTA SAT-1

Estaremos verificando cómo responde a esta carta.

Atentamente,


Lic. César Alfredo Lora Estrada
Gerente Regional Control
Superintendencia de Administración Tributaria
SAT

Base Legal: Artículos 98,112 y 146 del Código Tributario;
Artículo 3 de la Ley Orgánica de la Superintendencia de Administración Tributaria.



ALERTA SAT-4

Chimaltenango, Chimaltenango

Guatemala, 28 de mayo de 2014

Por favor presente su declaración del impuesto sobre la renta

Estimado contribuyente:

Hemos revisado nuestros archivos y encontramos que usted no ha presentado su declaración anual del Impuesto sobre la Renta correspondiente al año 2013.

Anteriormente, hemos considerado su falta como un descuido. Sin embargo, si usted no declara ahora, vamos a considerar que es su elección, y puede ser auditado y ser sujeto al procedimiento establecido por ley.

Por favor declare en un plazo de 10 días después de recibida esta alerta, lo cual será verificado.

- Para ello, visite: <http://declaraguatemala.gt> (Formulario SAT-1411)

Si usted no puede pagar el monto total ahora, podría ser posible que pague en cuotas, luego de preparar su declaración y acercarse a una oficina de la SAT.

Si tiene alguna consulta, contáctenos a través del teléfono 2329-7111. En ese caso haga referencia a su número de alerta: ALERTA SAT-4

Estaremos verificando cómo responde a esta carta.

Atentamente,

Lic. César Alfredo Laroj Estrada
Comisario Regional General de
Superintendencia de Administración Tributaria
SAT-4

Base Legal: Artículos 98, 112 y 146 del Código Tributario;
Artículo 3 de la Ley Orgánica de la Superintendencia de Administración Tributaria.

(T1) Original SAT Letter (n=6,198)



ALERTA SAT-1

Chimaltenango, Chimaltenango



Guatemala, 27 de mayo de 2014

Estimado contribuyente:

Usted es un ciudadano guatemalteco y Guatemala lo necesita. Sea un buen ciudadano y presente su declaración anual del Impuesto sobre la Renta del año 2013.

Por favor declare en un plazo no mayor a 10 días después de recibida esta alerta.

- Para ello, visite: <http://declaraguatemala.gt> (Formulario SAT-1411)

Si usted no puede pagar el monto total, podría ser posible que pague en cuotas, luego de preparar su declaración y acercarse a una oficina de la SAT.

Si tiene alguna consulta, contáctenos a través del teléfono 2329-7111. En ese caso haga referencia a su número de alerta: ALERTA SAT

¿Va a apoyar a su país?

Atentamente,



Lic. César Alfredo Lora Estrada
Gerente Regional Central
Superintendencia de Administración Tributaria

Base Legal: Artículos 98, 112 y 146 del Código Tributario;
Artículo 3 de la Ley Orgánica de la Superintendencia de Administración Tributaria.

2014 Taxpayer Choice Model (TCM): Designing Digital Communication Products To Reduce Phone and Mail Inventory

*Courtney Rasey and Mackenzie Wiley, Wage & Investment Research and Analysis,
Internal Revenue Service*

Introduction

Over the past decade, the proportion of American adults using the Internet has grown from 66 percent in 2005 to 87 percent in 2015.¹ Among Internet users, a majority (82 percent) have gone online in the past year to look for information or complete a transaction on a Government Website.² The Internal Revenue Service (IRS) Strategic Plan FY 2014–2017 cites the proliferation of Internet-based interactions as one of the major trends impacting the IRS currently and in the future. It states, “the growth of the Internet over the past decade has changed consumer expectations as they become increasingly more accustomed to using the Web for anything from ordering phone service to conducting transactions with financial institutions using traditional online and mobile devices. More and more customers show a preference for Internet-based service before trying other service channels such as phones, paper, or in-person.”³

As the population of Internet users grows and technology develops to provide more ways for individuals to get online, customer service channel offerings expand and IRS seeks to expand its customer service options, as well. While the IRS has made significant progress in digital service offerings (e.g., the Where’s My Refund? interactive tool and the Interactive Tax Assistant), the Strategic Plan acknowledges “there are distinct unmet taxpayer needs that provide opportunities for the IRS to introduce more online self-service options.”⁴ Even while facing severe budgetary constraints, IRS continues to plan and develop improved service channels to ease the burden for taxpayers trying to meet their tax obligations. While awareness of IRS service channels remains high among taxpayers (91 percent), the organization is committed to “expanding (its) portfolio of digital service offerings to meet customer expectations while continuing to keep taxpayer data secure.”^{5,6}

Goal I of the Strategic Plan is to “deliver high quality and timely service to reduce taxpayer burden and encourage voluntary compliance.”⁷ The following two Objectives of that goal emphasize consideration of new technologies to improve IRS’s customer service experiences for all taxpayers:

- Design tailored service approaches with a focus on digital customer service to meet taxpayer needs, preferences, and compliance behaviors in order to facilitate voluntary compliance.
- Provide timely assistance through a seamless, multichannel service environment to encourage taxpayers to meet their tax obligations and accurately resolve their issues.

¹ Pew Research Center, <http://www.pewinternet.org/data-trend/internet-use/latest-stats/>.

² Smith, Aaron. (2010, April 27). “Government Online,” Retrieved from Pew Research Center, <http://pewinternet.org/Reports/2010/Government-Online.aspx>.

³ Internal Revenue Service Strategic Plan FY 2014–2017. Retrieved from <http://www.irs.gov/pub/irs-pdf/p3744.pdf>.

⁴ Ibid.

⁵ “Internal Revenue Service Customer Satisfaction Survey—2014 W&I Taxpayer Experience Survey National Report,” Internal Revenue Service and Pacific Consulting Group, October 2014.

⁶ Internal Revenue Service Strategic Plan FY 2014–2017. Retrieved from <http://www.irs.gov/pub/irs-pdf/p3744.pdf>.

⁷ Ibid.

Strategizing and prioritizing the development of new service channels are difficult since research shows that individuals interact with Government through a mix of online and offline methods. Contacting the Government by phone and through online means are the most preferred channels among those surveyed in a 2010 study by the Pew Research Center.⁸ While the population as a whole prefers to contact Government agencies through the phone (35 percent) and then through online channels (28 percent), Internet users most prefer to contact Government agencies online (37 percent), with 33 percent reporting the phone as their second choice for service channel.⁹ The need for IRS services is great; the IRS 2014 Taxpayer Experience Survey shows that 41 percent of taxpayers have contacted the IRS for tax-related issue(s) in the last 12 months. Among those who contacted the IRS, the most-used service channel is the IRS Website (28 percent), followed by the IRS toll-free line (17 percent).¹⁰ For all but 1 of the 14 service needs, respondents were asked about in the survey, the IRS Website was most often indicated as taxpayers' first step toward resolving the issue (38 percent of those who needed to make a payment and 35 percent of those who needed information about a notice went to the Website first).¹¹ Additionally, there are many taxpayers who reported needing assistance with a tax-related issue in the past 12 months, but chose not to contact the IRS about it (46 percent).¹²

IRS cannot stop offering more costly traditional service channels, such as toll-free phone lines with a live customer service representative or local IRS offices, but it must determine how best to serve taxpayers with its limited resources. Digital communication products enable IRS to serve a greater number of customers at a lower cost, and IRS research indicates that taxpayers are ready for it. Some 56 percent of taxpayers said they would be likely to review their prior-year Federal tax return information online through a secure link on IRS.gov.¹³

A key strategy IRS employs to achieve its goals is the incorporation of taxpayer feedback into all aspects of its operations.¹⁴ The IRS uses multiple channels to obtain taxpayer feedback, including surveys, focus groups, interviews, comment cards and social media. The present research exemplifies the application of taxpayer feedback in IRS's strategic planning. Using conjoint-based models to prioritize potential new services, based on taxpayer preferences over traditional channels, helps the IRS tailor channel development to specific population segments and service needs. Designing and marketing digital communication products for taxpayers are critical components of IRS's strategic plan.

The purpose of this study is to determine how the IRS can best migrate taxpayers to less expensive service channels. To accomplish this overarching objective, findings prioritize potential digital communication products considered for development by IRS's Online Services (OLS) office.

This report begins by documenting the background and evolution of the IRS's conjoint-based research, and then focuses on a review of major findings from the 2014 survey effort and associated analysis resulting from its Taxpayer Choice Model.

Background

The 2014 effort to update the Taxpayer Choice Model (TCM) is the latest iteration of choice-based conjoint research dating back to 2005, when IRS first formed the Taxpayer Assistance Blueprint (TAB) team to address fundamental questions regarding the preferences and tax assistance needs of IRS customers. TAB was a large-scale study that impacted IRS business practices by strengthening its ability to make well-informed decisions about the way services are provided to its primary customers—individual taxpayers. It included a conjoint

⁸ Smith, Aaron. (2010, April 27). "Government Online," Retrieved from Pew Research Center, <http://pewinternet.org/Reports/2010/Government-Online.aspx>.

⁹ Ibid.

¹⁰ "Internal Revenue Service Customer Satisfaction Survey—2014 W&I Taxpayer Experience Survey National Report," Internal Revenue Service and Pacific Consulting Group, October 2014.

¹¹ Ibid.

¹² Ibid.

¹³ Ibid.

¹⁴ Internal Revenue Service Strategic Plan FY 2014–2017. Retrieved from <http://www.irs.gov/pub/irs-pdf/p3744.pdf>.

survey to investigate taxpayers' preference among the various options for contacting the IRS to resolve their tax-related issues.¹⁵

The 2006 conjoint study reviewed traditional service channels, such as Taxpayer Assistance Centers (TACs), toll-free phone line with help from a representative, toll-free phone line with automated menu options, Website browsing, and regular mail. The result of the 2006 conjoint study was the Taxpayer Value Model (TVM), an Excel-based model and predecessor to the TCM with an interface that enabled IRS to run "what if?" scenarios to understand how changes to service channel features affect taxpayer preference and value. Users of the tool are able to manipulate service channel attributes (such as access time, first contact resolution, etc.) to see how hypothetical changes in service impact taxpayers' likelihood of using a given channel for a tax-related task.

Since this service channel preference research was first conducted in 2006, the landscape of service channel offerings evolved and drastically changed. The IRS (more specifically, Wage and Investment Research & Analysis (WIRA)) continued the research, but with changes to reflect new service offerings.

In 2011, WIRA conducted another conjoint study to investigate taxpayers' preferences among options for contacting the IRS to resolve tax-related issues. The data collected in this research effort were used to develop a tool similar to the TVM that enabled IRS to run "what if?" scenarios; it is known as the Taxpayer Choice Model (TCM). The TCM, completed in 2011, helps the IRS understand how service channel changes affect taxpayer preference and value. The TCM was similar to the TVM, but featured newer service needs and service channel options to reflect changing trends in technology and taxpayer preference for digital options for tax-based interactions. In addition to asking respondents about their preferences for more established channels such as toll-free phone lines (with a customer service representative or automated menu prompts) and Website browsing, the survey gathered data about Web-based interactive tools, social media Websites, and smartphone applications.

In a follow-up study, WIRA worked with a team led by the Compliance function within IRS's Wage and Investment division to understand better taxpayer preference for digital communication options for taxpayers related to compliance tasks, such as those involved in completing a correspondence audit. The 2014 TCM includes service needs with corresponding service channels based on current availability and potential for future development.¹⁶

Objective

The main objective of this research is to determine how three service delivery dimensions—task, service channel, and channel performance—impact service channel selection among taxpayers. This comparative value analysis process evaluates service channel options for use by taxpayers.

Specifically, this project evaluates data for each of the service needs included in this study to accomplish the following objectives:

1. Understand taxpayer preference for current and potential service channels.
2. Determine how service channel changes affect taxpayer preference and taxpayer value.
3. Prioritize potential digital communication products based on how likely taxpayers prefer that service over current service channel options.
4. Identify the best ways to design and market digital communication tools in order to move taxpayers from high-cost service channels to both current and potential digital communication tools.

¹⁵ "Taxpayer Assistance Blueprint (TAB) Conjoint II Study Topline Report," Internal Revenue Service and Pacific Consulting Group, September 2006.

¹⁶ The six service needs included in the model are: Submit Documentation; Check Status of a Case; Sign a Document; Discuss Case Details; Set up a Payment Plan; and Request an Extension to Respond. Service channels varied by service need with the following service channels included for at least one service need: IRS Toll-Free Line, Live Assistor; IRS Toll-Free Line, Automated; IRS.gov Interactive Tool; Fax; Regular Mail; Local IRS Office; Smartphone Application; Automatic Email Communication; Automatic Text Notification; Secure Message; and Secure Online Chat.

Methodology

Conjoint Methodology

Choice-based conjoint analysis is a widely-used quantitative method in market research. It can be used to quantify customers' value of a specific product or service's attributes to determine which delivery scenario they most prefer among a set of choices.

In conjoint research focused on choosing a service channel, the researcher asks the respondents to assume they need to complete a specific task, such as getting information about the status of their refund. This provides respondents with realistic situations to evaluate and ultimately select their preference. Then, respondents see a series of scenarios showing varying requirement and outcome conditions. A requirement might be the taxpayer's provision of a Social Security number, which represents a certain level of burden. An outcome attribute might be the amount of time a task takes to complete (e.g., 5 minutes on the Web versus 30 minutes on the phone).

Respondents usually view between 10 and 30 conjoint scenarios with varying attribute levels and select the one they most prefer when comparing it with the others. Scenarios are designed using experimental design principles of independence and balance of features. By randomly varying the attributes shown to the respondents and observing responses to the scenarios, the researcher can statistically deduce what product or service features are more desired and which attributes have the most impact on choice. These realistic trade-off decisions provide more insight than simply asking respondents what they prefer or how important product and service features are to them.

The output of a choice-based conjoint survey is a set of preferences for each attribute in the study. In conjoint terminology, these preference sets are called part-worth utilities. Part-worth utilities can be used to summarize preferences of taxpayer segments and markets. Customer choices also can be projected by adding the part-worth utility for each individual attribute to calculate the overall preference for a specific service channel.

Since each respondent has a unique set of part-worth utilities, it is difficult and time-consuming for researchers to analyze the data by hand. Therefore, researchers typically create conjoint market simulators to access the data easily and quickly compare products or services. Using the simulator, researchers manipulate multiple attribute levels and even change the products or services considered in the model to create realistic what-if scenarios and see how customer preference changes in response to varied service delivery configurations.

Conjoint Survey Instrument Development

WIRA worked with a cross-functional workgroup to develop all aspects of the survey. The workgroup included the following IRS functions: Compliance, Online Services (OLS), and Business Modernization Office.

The team looked at volume data and the correspondence audit process to finalize the service needs in the survey instrument. Service channels for each service need were chosen based on current offerings, as well as those under consideration for development by Compliance and OLS.

Conjoint Design

WIRA used a fractional factorial design to select the attribute levels for each respondent. Fractional factorial design creates profiles to limit the number of combinations the respondent evaluates while ensuring enough data are available for statistical analysis. This results in a carefully controlled set of "profiles" for the scenarios. Each respondent received five scenarios for each of two service needs (ten scenarios in total).

Service Needs and Service Channels

The service channels for each service need were chosen based on current offerings and service channels Compliance and OLS are considering for development (see Table 1).

TABLE 1. Service Channels Included by Service Need

Service Needs	Service Channels
1. Submit Documentation	<ol style="list-style-type: none"> 1. Toll-Free Phone, Live Assistor 2. Fax 3. IRS Website—Interactive Tool 4. Smartphone Application 5. Regular Mail 6. Secure Message 7. Secure Online Chat
2. Get the Status of a Case/Transaction	<ol style="list-style-type: none"> 1. Toll-Free Phone, Live Assistor 2. Toll-Free Phone, Automated 3. IRS Website – Interactive Tool 4. Smartphone Application 5. Automatic Email Communication 6. Automatic Text Notification 7. Secure Online Chat
3. Sign a Document	<ol style="list-style-type: none"> 1. Toll-Free Phone, Live Assistor 2. Fax 3. IRS Website—Interactive Tool 4. Smartphone Application 5. Regular Mail 6. Secure Message 7. Secure Online Chat
4. Get Information About a Notice You Received/ Discuss Case Details	<ol style="list-style-type: none"> 1. Toll-Free Phone, Live Assistor 2. Regular Mail 3. Secure Message 4. Secure Online Chat
5. Set Up a Payment Plan	<ol style="list-style-type: none"> 1. Toll-Free Phone, Live Assistor 2. IRS Website—Interactive Tool 3. Regular Mail 4. Local IRS Office
6. Request an Extension	<ol style="list-style-type: none"> 1. Toll-Free Phone, Live Assistor 2. Toll-Free Phone, Automated 3. IRS Website—Interactive Tool 4. Regular Mail 5. Smartphone Application

Attributes

The team chose to include the following attributes (also referred to as service channel features) in the survey instrument:

- **Time Required** (how long it takes to receive the service, which includes wait time and service time);
- **Confirmation of Receipt** (whether receive confirmation of the transaction and/or response details);
- **Authentication** (how taxpayers prove that they are who they say they are in order to use the service);
- **Account Required** (whether the taxpayer must create an account in order to use the service); and
- **Account Update Time** (amount of time between service and transaction recorded on the tax account).¹⁷

Survey Participant Recruitment

WIRA administered the survey through a contractor, Fors Marsh Group (FMG). Participants were randomly selected from a panel owned by one of FMG's subcontractors.

¹⁷ The survey instrument included Account Update Time only for the service need "Submit Documentation."

The subcontractor invited potential respondents to participate in a survey on their preferences for help with tax-related service needs. Participants were provided a link/Web address to a secure Website and a unique PIN to access and respond to survey questions. Respondents accessed the Website and answered the preliminary questions pertaining to demographics and their history of interactions with the IRS. These included screening questions aimed at determining whether the respondent qualified for the survey based on predetermined criteria.

For this study, the screening criteria ensured that each participant:

- was 18 years of age or older;
- filed a tax return for the most recent tax year; and
- was not a current IRS employee.

Survey Administration

The survey was administered in May 2014 and WIRA received the final data from the contractor in July. A total of 2,303 respondents completed the survey with each seeing two scenarios. Sample size for each service need ranged from 766 to 770 respondents.

Weighting

WIRA created weights for each participant based on age, tax return filing method, tax return preparation method, gender, geographic region, and household income. Analysts used Compliance Data Warehouse (CDW) historical tax return filing data to create the weights for age, filing method, and preparation method, and Current Population Survey (CPS) for gender, region, and household income.

Data Analysis

WIRA conducted Hierarchical Bayes analysis using Sawtooth Software to estimate part-worth utilities, and ultimately used this analysis to create the 2014 Taxpayer Choice Model (TCM). Analysts used the 2014 TCM to analyze taxpayer preference employing scenarios developed with OLS.

Limitations

Conjoint analysis is a quantitative research method used to predict future behavior based on specific service channel configurations. This prediction is based only on the services and service channel features included in the study. When designing a conjoint study, researchers carefully consider what service channel features most influence service channel choice and should be included. However, other service channel features not included in the study can affect taxpayer choice. In addition, conjoint analysis assumes perfect awareness and access to all service channels, which does not hold true for all taxpayers. Due to these caveats, conjoint results serve to show the direction and magnitude of a preference shift, not an exact numeric estimate. Results should be used in conjunction with other data when making operational decisions.

As with all quasi-experimental designs, threats to internal and external validity exist. Potential threats to internal validity include local historical effects and testing effects. Local history can also threaten external validity as contexts for decisions can change based on context.

Key Findings

Taxpayer Preference and Value

WIRA studied taxpayer preference and migration to potential service channels by service need. Service channels for each service need were chosen based on current offerings and service channels that Compliance and OLS are considering for future development. WIRA researchers analyzed operational data and worked with IRS subject-matter experts to design the scenarios.

This section of the report explains the findings from scenarios run in the TCM for each of the six service tasks listed in Table 1, and correspondingly is divided into subsections, listed 1 through 6 below. Each service task was given a base case that accurately reflects the current state of IRS service offerings (i.e., available service channels and their corresponding performance attribute levels) and then a test case scenario where researchers simulated the features of future state IRS service channels.

“Future state” scenarios can vary. The tool enables researchers to look at how taxpayer preference for channels shifts due to changes in service channel attributes (such as adjusting the time it takes to complete a given task or changing the amount of information required from a taxpayer to authenticate their identity), as well as the addition or removal of service channel options to complete a given task.

Generally, for each service task covered in this section, the accompanying graphical figure shows two states (Figure 1). The first shows taxpayer preference for current and potential service channels. The second shows taxpayer preference changes under a new hypothetical service delivery configuration.

1. Submit Documentation

In the current state, taxpayers have three options to submit documentation to the IRS for various issues and needs: calling (phone (*Customer Service Representative (CSR)*)), writing (*Regular Mail*), or faxing (*Fax*). The IRS is considering two additional digital options in the future for this task: *Secure Message* and *Secure Online Chat*. Using the TCM to simulate a scenario for the potential new service channels and their attributes, results show that taxpayer preference shifts away from the traditional channels to new, digital options. Taxpayers still have a high preference for contacting the IRS via the phone and speaking to a customer service representative (27 percent); but they are willing to submit documentation through secure digital options (35 percent).

FIGURE 1. Current and Future State Taxpayer Preference for Submit Documentation

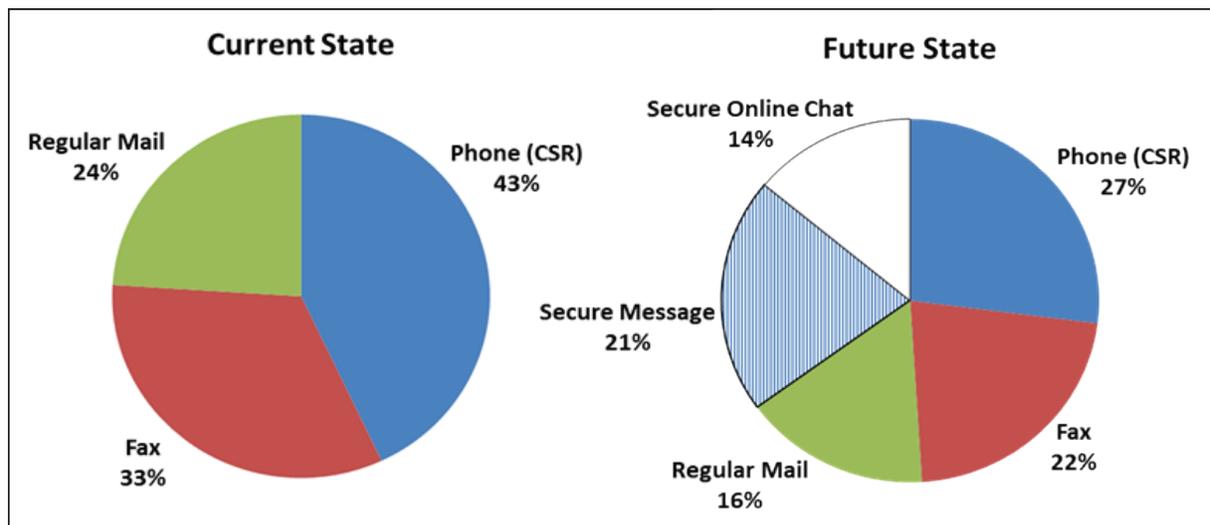
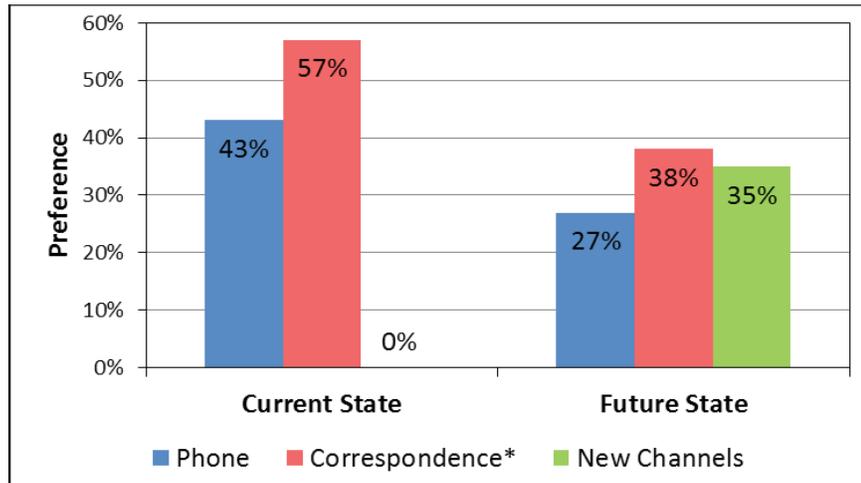


Figure 2 demonstrates potential migration from the phone and correspondence stream to the new channels that IRS may offer in the future. *Secure Message* and *Secure Online Chat* shift 35 percent of the preference for traditional channels, with 16 percent shifting away from *Phone (CSR)* and a 19-percent decrease in the correspondence stream (8 percent *Regular Mail* decrease and 11 percent *Fax* decrease).

FIGURE 2. Taxpayer Migration from Traditional Channels for Submit Documentation



* Information sent via Fax is considered part of the correspondence stream.

2. Get the Status of a Case/Transaction

Taxpayers can get the status of their case in the current state only by calling the IRS.¹⁸ The IRS is considering six potential service channels for the future state. Figure 3 illustrates taxpayer preference in this potential future state. With phone as the only current option, current state preference is considered 100 percent for *Phone (CSR)*. TCM results indicate taxpayers have a strong preference to receive updates about the status of their case through *Automatic Text Notification* (27 percent) or through an *Online Tool* (19 percent) available on IRS.gov.

FIGURE 3. Future State Taxpayer Preference for Get the Status of a Case

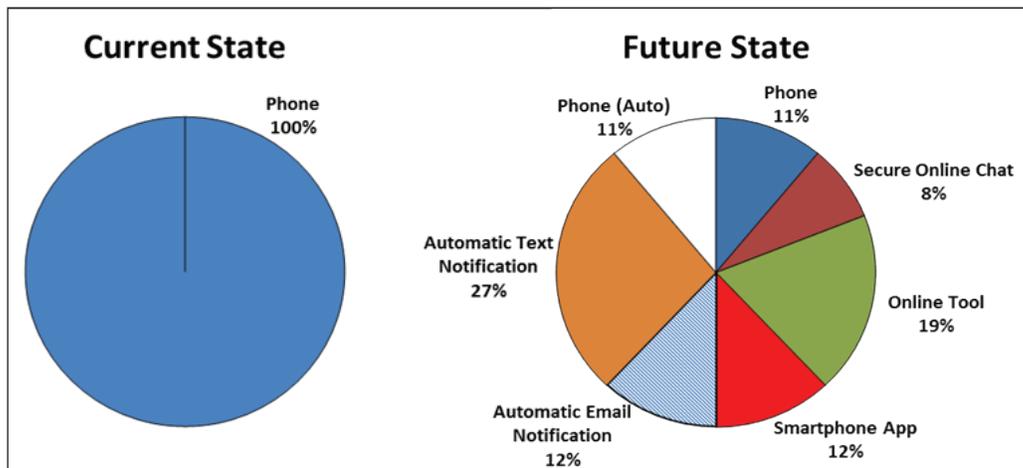
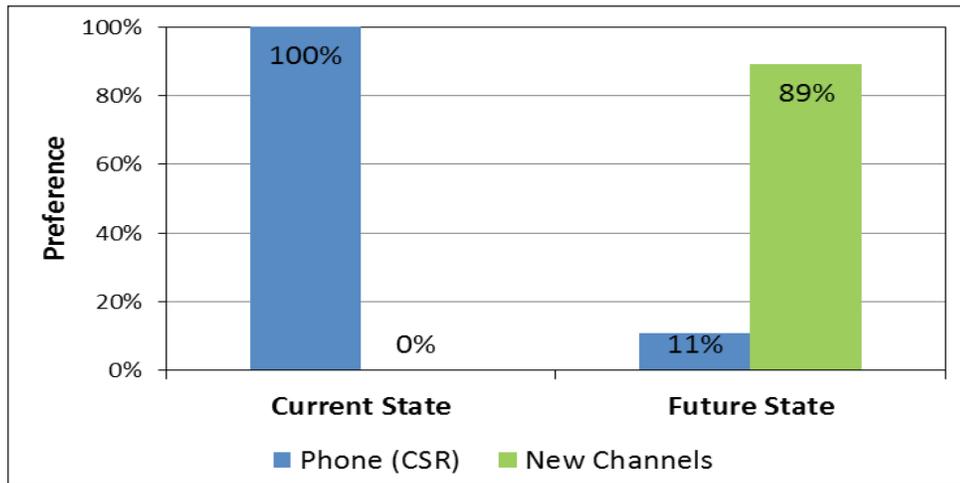


Figure 4 demonstrates the shift away from *Phone (CSR)* due to the addition of the new service channels. Most taxpayers prefer to use a new channel in the future state, with an 89 percent migration away from the phone.

¹⁸ While taxpayers could choose to visit a local IRS office or write a letter to get the status of their case, those channels are not included in the current state of this model due to low volume.

FIGURE 4. Taxpayer Migration from Traditional Channels for Get the Status of a Case



3. Sign a Document

In the current state, taxpayers can write, call or fax the IRS to sign a document.¹⁹ The IRS may offer *Secure Message* and an *Online Tool* in the future for this task. Figure 5 illustrates TCM scenario results showing taxpayer preference in the current state and the potential future state. The TCM shows that while some taxpayers’ preference for traditional channels will not change, many have a strong preference for being able to sign a document via online channels, such as a *Secure Message* exchange with the IRS (38 percent) or an *Online Tool* on IRS.gov (18 percent).

FIGURE 5. Current and Future State Taxpayer Preference for Sign a Document

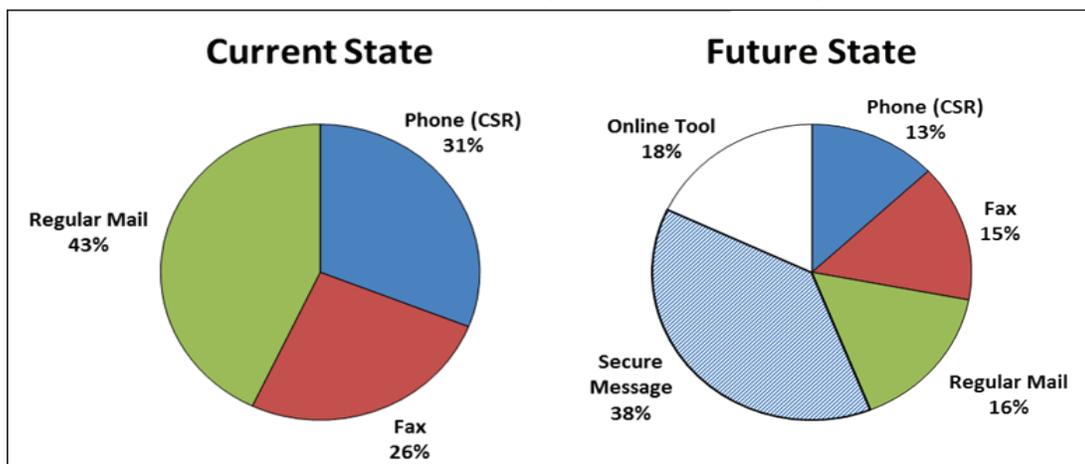
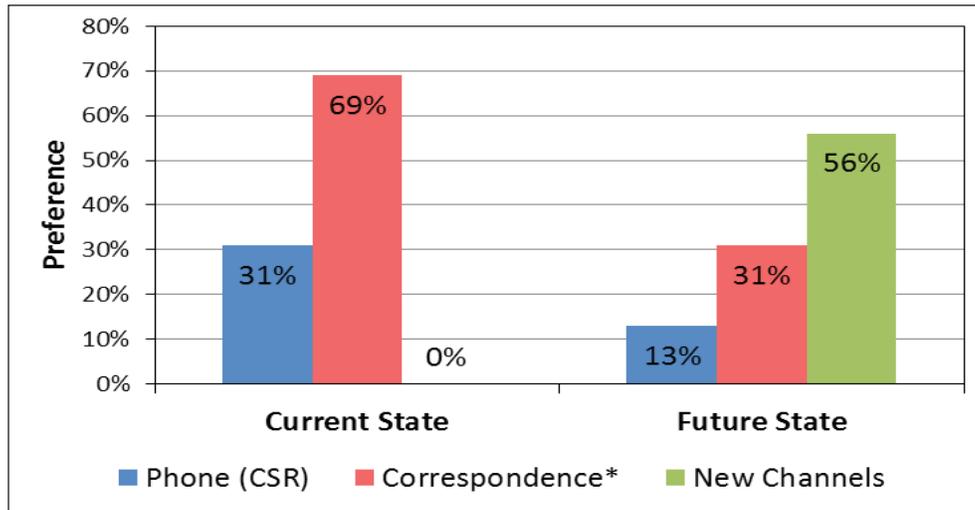


Figure 6 demonstrates the shift away from phone and correspondence streams due to the addition of the new service channels. Those new channels shift 56 percent of the preference for traditional channels, with an 18-percent shift away from *Phone (CSR)* and 38-percent decrease in the correspondence stream (27 percent *Regular Mail* decrease and 11 percent *Fax* decrease).

¹⁹ If taxpayers call the IRS to sign a document, they are able to resolve their issue using an internal system to fax a document to the customer service representative during their interaction.

FIGURE 6. Taxpayer Migration from Traditional Channels for Sign a Document



* Information sent via Fax is considered part of the correspondence stream.

4. Discuss Case Details

In the current state, taxpayers can write or call the IRS to discuss case details. The IRS may offer *Secure Message* and *Secure Online Chat* in the future for this task. Figure 7 illustrates taxpayer preference in the current state and this potential future state. While addition of *Secure Message* or *Secure Online Chat* as service channel options would shift taxpayer preference from the traditional channels of *Phone (CSR)* and *Regular Mail*, the shift to digital service options is not as strong as with other service needs considered in the TCM and described in this paper.

FIGURE 7. Current and Future State Taxpayer Preference for Discuss Case Details

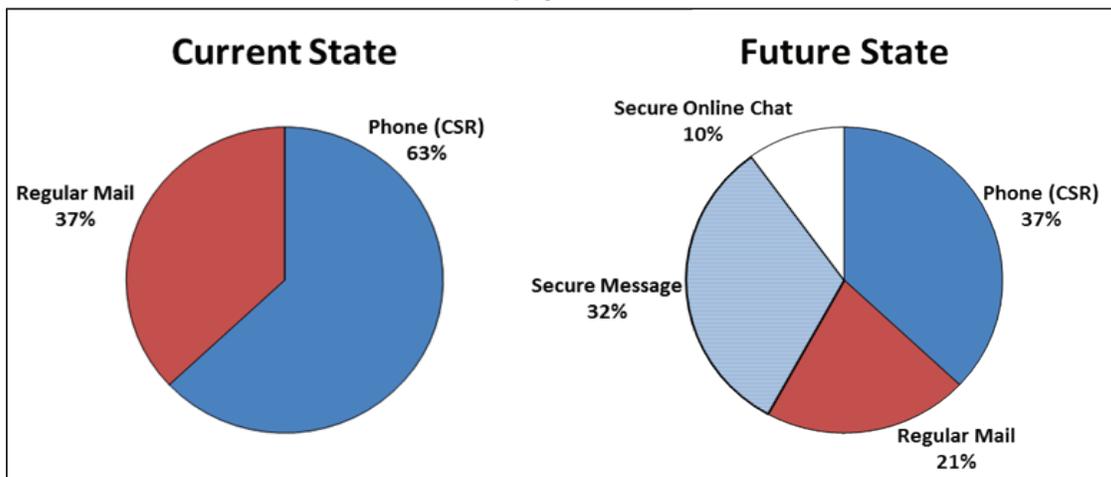
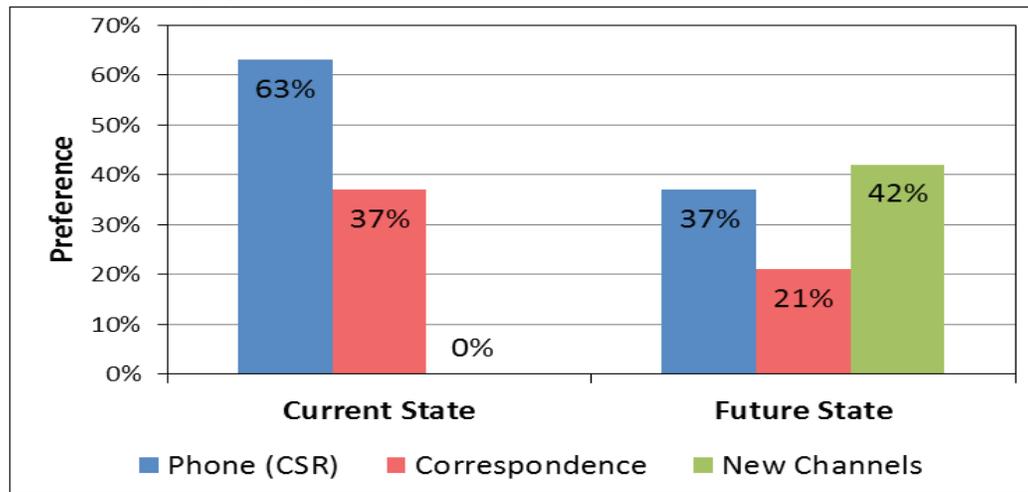


Figure 8 demonstrates the shift away from the phone and mail from the addition of the new service channels. *Secure Message* and *Secure Online Chat* shift 42 percent of the preference for traditional channels, with a 26-percent shift away from *Phone (CSR)* and 16-percent decrease in the correspondence stream via *Regular Mail*.

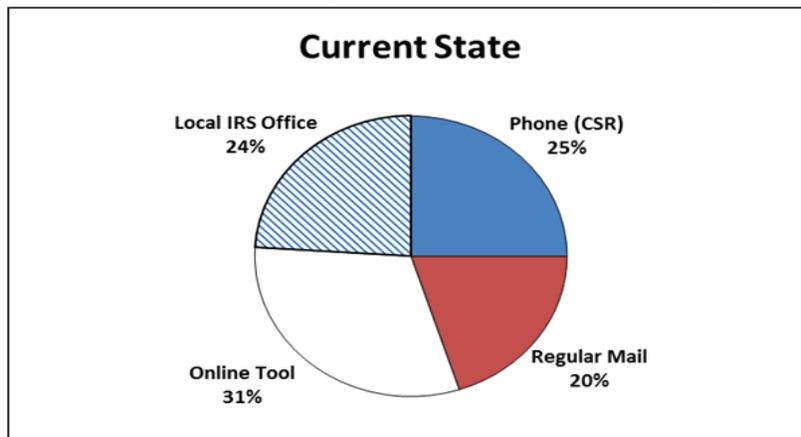
FIGURE 8. Taxpayer Migration from Traditional Channels for Discuss Case Details



5. Set Up a Payment Plan

In the current state, taxpayers can write the IRS, call the IRS, go to a local IRS office, or use an online tool to set up a payment plan. IRS subject-matter experts and conjoint team members indicated no current plans for new payment plan offerings. As such, Figure 9 illustrates taxpayer preference in the current state.

FIGURE 9. Current Taxpayer Preference for Set Up a Payment Plan

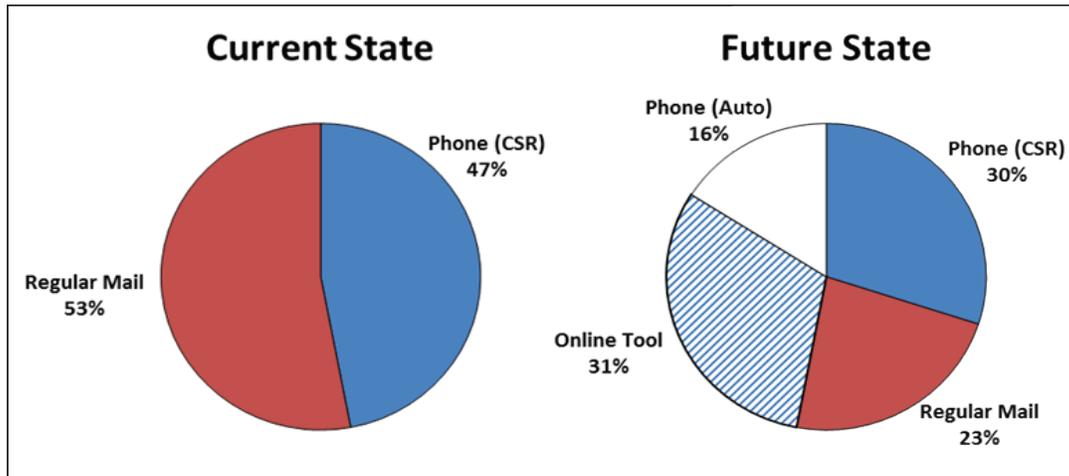


While this report displays only the current taxpayer preference for this service need, the TCM is versatile enough to allow researchers to run scenarios to understand taxpayer preference among various market segments if changes were made to the features of any of the four current service channels. Examples of the demographic analysis of market segments can be found later in this report.

6. Request an Extension

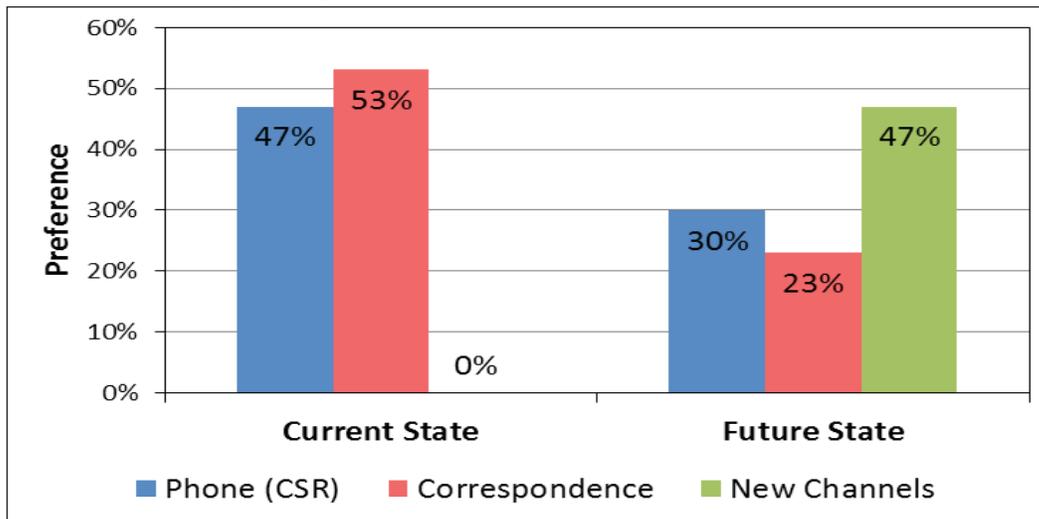
In the current state, taxpayers can write or call the IRS to request an extension to respond. The IRS may offer an *Online Tool* or *Automated Phone Line* in the future for this task. Figure 10 illustrates taxpayer preference in the current state and this potential future state.

FIGURE 10. Current and Future State Taxpayer Preference for Request an Extension



While more than half of taxpayers (53 percent) still prefer traditional channels in the future state, Figure 11 demonstrates the shift away from the phone and mail when offered the new service channels. *Online Tool* and *Automated Phone Line* shift 47 percent of the preference for traditional channels, with a 17-percent shift away from *Phone (CSR)* and a 30-percent decrease in the correspondence stream via *Regular Mail*.

FIGURE 11. Taxpayer Migration from Traditional Channels for Request an Extension



Prioritize Potential Digital Communication Products

WIRA examined taxpayer preference for potential stand-alone offerings by calculating the effect of adding only one new service channel rather than the entire future state as outlined above.

Table 2A provides the percentage of preference shift predicted from adding *one* particular new service for a given service need. “Stand-alone service” is defined as a new service channel aimed at accomplishing a specific task, such as requesting an extension. The TCM does not account for demand for service need; therefore, the table represents only the preference shift for taxpayers. Before prioritizing possible future service channel offerings, IRS would need to consider this preference data in combination with other internal data sources.

Table 2B provides the same information, but in declining order of the magnitude of the preference shift. The new stand-alone services for Get the Status of a Case have the highest preference shift away from the traditional channels. The top six stand-alone services in Table 2B are all for Get the Status of a Case, and *Automatic Text Notification* has the highest preference shift overall.

TABLE 2A. Stand-Alone Service Shift by Task and New Service

Task	Stand-Alone Service Channel	Preference Shift to Service (as stand-alone addition)
Submit Documentation	<i>Secure Message*</i>	31%
	<i>Secure Online Chat</i>	25%
Get the Status of a Case	<i>Automatic Text Notification*</i>	62%
	<i>Toll-Free Line, Automated</i>	59%
	<i>Smartphone App</i>	58%
	<i>Online Chat</i>	50%
	<i>Online Tool</i>	48%
	<i>Automatic Email Notification</i>	48%
Sign a Document	<i>Secure Message*</i>	43%
	<i>Online Tool</i>	33%
Discuss Case Details	<i>Secure Message*</i>	39%
	<i>Secure Online Chat</i>	30%
Request an Extension	<i>Online Tool*</i>	36%
	<i>Smartphone App</i>	23%
	<i>Toll-Free Line, Automated</i>	23%

* These service channels show the highest preference shift as a stand-alone service channel addition for that service need.

TABLE 2B. Stand-Alone Service Shifts by Strength of Preference

Stand-Alone Service Channel	Task	Preference Shift to Service (as stand-alone addition)
<i>Automatic Text Notification*</i>	<i>Get the Status of a Case</i>	62%
<i>Toll-Free Line, Automated</i>	<i>Get the Status of a Case</i>	59%
<i>Smartphone App</i>	<i>Get the Status of a Case</i>	58%
<i>Online Chat</i>	<i>Get the Status of a Case</i>	50%
<i>Online Tool</i>	<i>Get the Status of a Case</i>	48%
<i>Automatic Email Notification</i>	<i>Get the Status of a Case</i>	48%
<i>Secure Message*</i>	<i>Sign a Document</i>	43%
<i>Secure Message*</i>	<i>Discuss Case Details</i>	39%
<i>Online Tool*</i>	<i>Request an Extension</i>	36%
<i>Online Tool</i>	<i>Sign a Document</i>	33%
<i>Secure Message*</i>	<i>Submit Documentation</i>	31%
<i>Secure Online Chat</i>	<i>Discuss Case Details</i>	30%
<i>Secure Online Chat</i>	<i>Submit Documentation</i>	25%
<i>Smartphone App</i>	<i>Request an Extension</i>	23%
<i>Toll-Free Line, Automated</i>	<i>Request an Extension</i>	23%

* These service channels show the highest preference shift as a stand-alone service channel addition for that service need.

Together, Table 3, Table 4, and Table 5 illustrate that *Secure Message* has the highest preference shift for all service needs for which it is being considered as a potential option, even over other digital service channels. The service needs for which *Secure Message* is an option in the TCM are Sign a Document (Table 3), Discuss Case Details (Table 4), and Submit Documentation (Table 5).

TABLE 3. Stand-Alone Channel Ranking for Secure Message (Sign a Document)

Service Channel for Sign a Document	Preference Shift to Channel (as stand-alone addition)
Secure Message	43%
Online Tool	33%

TABLE 4. Stand-Alone Channel Ranking for Secure Message (Discuss Case Details)

Service Channel for Discuss Case Details	Preference Shift to Channel (as stand-alone addition)
Secure Message	39%
Secure Online Chat	30%

TABLE 5. Stand-Alone Channel Ranking for Secure Message (Submit Documentation)

Service Channel for Submit Documentation	Preference Shift to Channel (as stand-alone addition)
Secure Message	31%
Secure Online Chat	25%

Demographic Analysis

The demographic analysis included the following demographic categories:

- Age
- Tax return preparation method
- Previous channel use
- Previous post-filing contact

In general, taxpayers who have used digital communication products outside of IRS interactions are more likely to use those channels with the IRS for all service needs than those who have not.

WIRA identified additional demographic differences for some of the service needs included in the study. However, analysis did not show significant differences between demographic segments for Discuss Case Details and Request an Extension. Demographic analysis for the remaining service needs (Submit Documentation, Get the Status of a Case, and Sign a Document) is included below.²⁰

For the *Submit Documentation* service need, there were some noteworthy shifts from traditional channels among a few market segments compared to the general population. Table 6 and Table 7 show the demographic differences WIRA discovered. Table 6 lists the demographic segments with a higher shift from traditional channels compared to the general population, defined as having at least 40-percent shift away from traditional channels to the new service channels.

²⁰ There is no additional demographic analysis included for *Set Up a Payment Plan* because there are no plans at this time to expand service channel offerings for that service need.

TABLE 6. Demographic Segments with At Least a 40 Percent Preference Shift from Traditional Channels (compared to general population) for *Submit Documentation*

Demographic Segment	Percentage Preference Shift
36-45 years old	40%
Self-prepared return using tax software/Free File	40%
Used secure message system previously	47%
Used online chat previously	40%
General population	35%

Table 7 lists the four demographic segments with at least 25-percent shift to *Secure Message*, specifically. (*Secure Message* was the channel with the greatest potential as a stand-alone new service, as shown in Table 2B.) Three of the four segments are the same ones listed in Table 6, indicating that at least 25 percent of the high preference shift from traditional channels for those segments was due to a shift to *Secure Message*.

TABLE 7. Demographic Segments with At Least a 25 Percent Preference Shift to *Secure Message* (compared to general population) for *Submit Documentation*

Demographic Segment	Percentage Preference Shift
18-24 years old	27%
Self-prepared return using tax software/Free File	25%
Used secure message system previously	26%
Used online chat previously	26%
General population	21%

For Get the Status of a Case, the overall shift away from Phone (CSR) (the only option currently available to taxpayers for this task) is similar for all demographic groups; 89 percent of all taxpayers and a range of 81-92 percent of taxpayers among each of the various demographic segments included in the TCM prefer another service channel offering. The demographic segments with the highest shift to *Automatic Text Notification* (the channel with the greatest potential shift as a stand-alone new service) are those with post-filing contact (41 percent preference shift), 18-24 years old (36 percent preference shift), and 36-45 years old (34 percent preference shift).

For Sign a Document, taxpayers aged 18-24 years old and 66-75 years old had the highest preference shift to an *Online Tool* (23 percent of each demographic segment prefer an *Online Tool*). Those who had used on-line chat previously and taxpayers aged 36-45 years old had the highest preference shift to *Secure Message* (48 percent of each demographic segment prefer *Secure Message*).

Attribute Importance

Attribute importance is defined as how much difference each service channel feature makes in the total preference for a product or service. It shows the “power” of each attribute in changing service channel preference/choice. When calculations show the service channel attribute has the largest impact on a service channel choice, ultimately, a portion of a taxpayers’ choice cannot be changed by adjusting service channel features. However, adjusting service channel attributes in the following ways can have an impact on taxpayers’ choice of service channel:

- increasing or decreasing the time it takes to complete a task,
- enabling the service channel to provide the channel user a confirmation of receipt for the task,

- changing the amount and sensitivity of information required for authentication, and
- requiring individuals to create an account in order to use a service channel.²¹

Attribute importance results for the 2014 TCM service needs are included below.²²

Figure 12 illustrates attribute importance for Get the Status of a Case and Request an Extension; results show that they have the same service channel feature order of importance. Service channel has the largest impact on service channel choice, which means a portion of a taxpayer's choice cannot be changed by adjusting service channel features. The service channel features have the following order of importance (listed in highest-to-lowest order): ID Proofing, Confirmation of Receipt, Account Required, and Time Required.

FIGURE 12. Attribute Importance for Get Status of a Case and Request an Extension

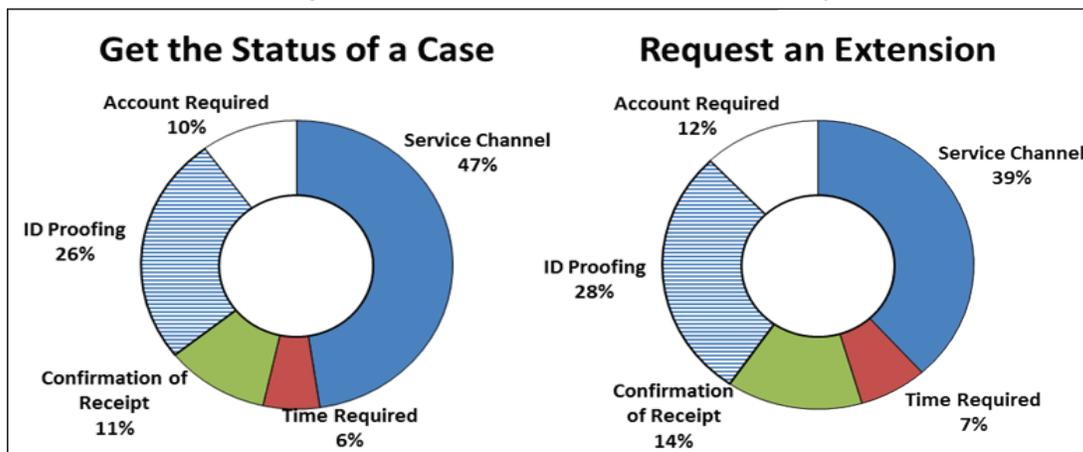
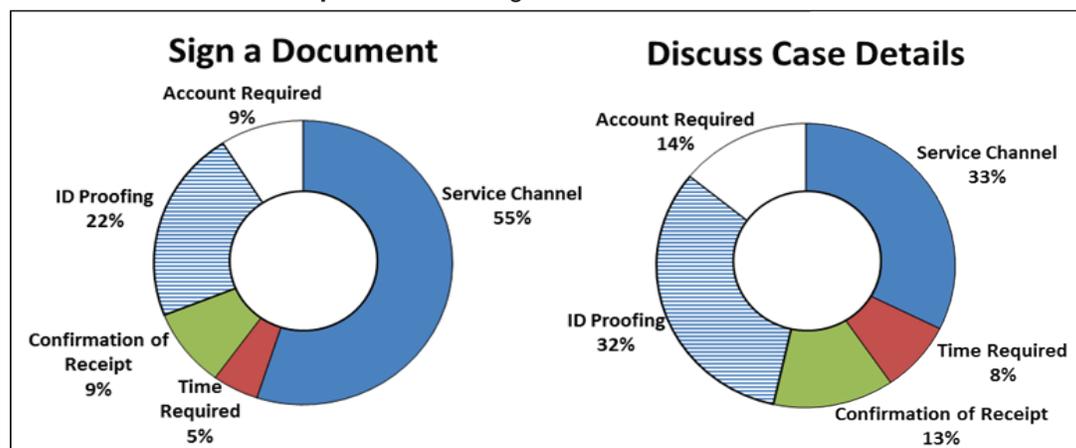


Figure 13 illustrates attribute importance for Sign a Document and Discuss Case Details as they have the same service channel feature order of importance. Service channel has the largest impact on service channel choice for both service needs, with service channel accounting for 55 percent of a person's choice for Sign a Document and 32 percent for Discuss Case Details. The service channel features have the following order of importance (listed in highest-to-lowest order): ID Proofing, Account Required, Confirmation of Receipt, and Time Required.

FIGURE 13. Attribute Importance for Sign a Document and Discuss Case Details

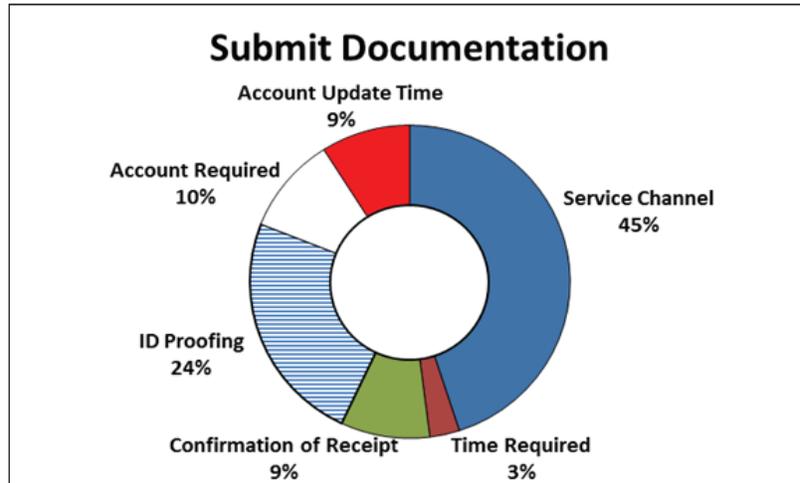


²¹ For the service need *Submit Documentation*, WIRA also looked at the amount of time between the time of service and the transaction being recorded on an individual's tax account.

²² There is no attribute importance information included for *Set Up a Payment Plan* because there are no plans at this time to change service channel features for that service need.

Figure 14 illustrates attribute importance for Submit Documentation. Service channel has the largest impact on service channel choice, with 45 percent. Submit Documentation has a fifth service channel feature included, Account Update Time.

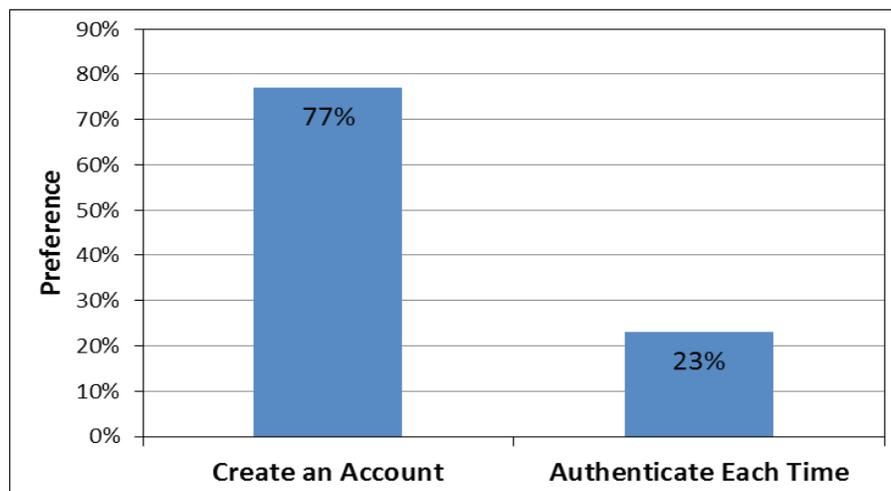
FIGURE 14. Attribute Importance for Submit Documentation



Authentication

Respondents were asked directly whether they would prefer to create an account or authenticate their identity each time they needed to use an IRS service. Figure 15 shows that the majority of respondents (77 percent) indicate that they would prefer to create an account that allows them to sign in using a username and password for subsequent service needs.

FIGURE 15. Authentication Preference Percentage



Conclusions

WIRA draws the following conclusions related to new product creation, product features, and marketing new products. These conclusions and recommendations are based on TCM results and other WIRA research. Other information, such as development costs and technological barriers, must be considered along with these recommendations.

New Product Creation

Creating alternative service channels for getting the status of a case or transaction could potentially reduce taxpayer calls to the IRS. Currently all contacts must go through the phone and CSR, but TCM results show that many taxpayers prefer to contact IRS a different way for that task. Creating even one additional service channel could reduce the high volume of contacts for this task by shifting taxpayers to the new channel. W&I Reporting and Compliance Annual Satisfaction Survey results support developing alternative channels for updating taxpayers on the status of their case (FY12 and FY13 reports).

Secure messaging has the potential to shift preference from traditional channels. The TCM shows a high preference shift to secure message as a stand-alone tool for all service needs where secure message is an option.

The IRS should investigate creating an online tool for requesting an extension to respond to a notice. Online tool shifts the most preference away from phone and mail for this task.

Product Features

Providing taxpayers with a confirmation of receipt for digital communication products affords greater resolution for taxpayers while they wait for their account to be updated or to receive a response. Both the 2007 and 2011 conjoint studies WIRA conducted indicate that resolution and degree of progress towards resolution affect taxpayer channel choice and are important to taxpayers. It also could potentially reduce contacts related to getting the status of a case by providing a generic timeline.

Creating digital self-service options for taxpayers via an online account could potentially reduce phone and mail traffic to the IRS. Most respondents reported a preference for creating an account that allows them to authenticate in less time for subsequent service needs.

Marketing New Products

Based on the demographic analysis, WIRA recommends partnering with tax return preparation software companies and Free File sites to increase awareness of new digital communication products. In addition, the IRS should consider using outgoing correspondence (i.e., notices) and/or their inserts to increase awareness of automatic text notification of status changes.

Appendix

Service Channel and Service Need Descriptions Provided to Survey Respondents

Toll-Free Phone Line

One method of help is the IRS Toll-Free phone line, which offers an automated self-help menu and access to live phone assistors.

IRS Website

Next, let's talk about the IRS Website that can be reached at www.irs.gov. You must have Internet access to use the IRS Website. Once at the Website, you can browse through different pages to find information about filing taxes or use interactive tools. Interactive tools are electronic tools that help you perform tasks on irs.gov, such as calculating withholding amounts, determining your filing status, or finding out the status of your refund.

Smartphone

The IRS has begun offering applications for smartphones. A smartphone is a device that a person can use to make telephone calls, but also adds in features that you might find on a personal digital assistant or a computer. Examples of smartphones include the iPhone and Android smartphones (i.e., HTC Evo, Motorola Droid X, Samsung Galaxy, etc.).

Individuals can download applications for smartphones that allow them to complete tasks or get information. An example of a current smartphone application offered by the IRS is IRS2Go. IRS2Go lets taxpayers check on the status of their tax refund and obtain other helpful tax information.

Regular Mail

The IRS also provides service through regular mail. Taxpayers can use the regular mail to ask tax questions and send payments. Once the IRS receives mail from a taxpayer, it typically takes between 30 and 90 days for the taxpayer to receive a response.

Fax

Taxpayers can transfer documents to the IRS via fax. For example, taxpayers can fax receipts or signed documents to the IRS.

Secure Message

The IRS may offer e-mail communication to taxpayers through a secure portal similar to secure message systems used by some health insurance companies, credit card companies, banks, and loan institutions. In this type of communication, people can log onto the company's Website and communicate with the company by sending and receiving messages. In some instances, customers may attach digital documents to the message.

For most tasks, it would take between 30 and 90 days for the taxpayer to receive a response back from the IRS.

Automatic Notification

The IRS may offer automatic e-mail or text notifications for account status updates similar to account notifications provided by some banks and credit card companies to inform customers of low balances, payments processed, or payments due.

For the automatic e-mail notifications, taxpayers would sign up for the service and then get emails whenever a change to their account occurs, such as a payment posting to their account.

For the automatic text notifications, taxpayers would sign up for the service online and then get texts whenever a change to their account occurs, such as a payment posting to their account.

Local IRS Office

The IRS runs local IRS offices in many cities throughout the country. Locations vary with most people living about 30 minutes from an IRS office. At these offices, you can walk in and get a number of services related to paying your taxes. You can get forms, get questions answered, make a payment, or set up a payment plan. IRS service representatives are available to help you in person, if necessary.

Secure Online Chat

The IRS may offer secure online chat through the irs.gov Website where taxpayers can communicate with a live assistor through the Internet by typing messages to one another in real time. Some insurance companies, banks, and other companies have this feature on their Webpage for individuals who cannot find answers to their questions through browsing the Web.

Service Need Descriptions from Survey Instrument

People need help with many issues related to taxes. For each of the service needs listed, indicate if you have ever needed help to complete the task.

- (1) Submit documentation to the IRS to verify items listed on your tax return. For example, document the itemized deductions you listed on your Schedule A, provide receipts for your medical expenses, or provide receipts for charitable donations.
- (2) Get the status of your case. For example, find out if the IRS received a payment or documentation you sent.
- (3) Sign a document. For example, sign a document to show you agree with the proposed changes to your tax return.
- (4) Ask for additional time to respond to a notice you received from the IRS.
- (5) Get more information about a notice you received from the IRS. For example, get more information about what documents the IRS needs you to send or why the documents are needed.
- (6) Make a payment or set up a payment plan.

Examining and Addressing Taxpayer Expectations for Affordable Care Act (ACA) Automated Information

Ariel S. Wooten and Marisa E. McDaniels,

Wage & Investment Research and Analysis, Internal Revenue Service

Introduction

The Patient Protection and Affordable Care Act (PPACA), commonly referred to as the Affordable Care Act (ACA), is a United States (U.S.) Federal statute signed into law by President Barack Obama on March 23, 2010. Together with the Health Care and Education Reconciliation Act, it represents the most significant regulatory overhaul of the U.S. healthcare system since the passage of Medicare and Medicaid in 1965.¹ The ACA has 500 provisions, with more than 40 provisions amending or adding to the existing tax code. The applied changes increase the complexity of tax administration, the administrative workload of the Internal Revenue Service (IRS), and the level of burden on taxpayers. Examples of some of the applied changes, additional IRS duties, and increased burdens on taxpayers include:

- Collecting new types of information from employers, insurers, and taxpayers: The ACA requires employers to tell the IRS whether an employee has health insurance. At the end of the year, insurers must provide the IRS and policyholders a form verifying coverage status, and individuals must include those forms with their annual Federal tax return. Taxpayers must be sure that they report coverage accurately on their return, and that their insurance coverage is adequate, based on the ACA coverage requirements.
- Determining who qualifies for subsidies, Medicaid, or exemptions: The ACA will provide Federal subsidies (known as premium tax credits) to help millions of people pay for their health insurance coverage. When an individual applies for coverage on an exchange (either Federal or State), the information provided will be cross-checked with income, job, and coverage information from the IRS to determine eligibility for a premium credit or Medicaid. Taxpayers who meet certain criteria may be eligible for a coverage exemption or a premium tax credit. Outside of traditional Federal tax return filing protocol, additional steps must be taken by taxpayers to apply, qualify, and calculate exemption or credit amounts.

Administration of ACA has provided the IRS with unique challenges and opportunities for success. In order to assist taxpayers with voluntary compliance by addressing the vast array of preliminary taxpayer questions, the IRS developed an automated phone line to answer common ACA-related questions.

The IRS recently explored individual and small business taxpayer expectations for ACA information that is provided via the automated phone line. This research asked participants to work through testing scenarios using the automated toll-free line, answer questions gauging their comprehension of the material presented on the line, rate their overall experience, and participate in a focus group discussion regarding their expectations and suggestions for improvement.

¹ Information provided by the Internal Revenue Service (IRS) Joint Operations Center (JOC) Affordable Care Act (ACA) liaison, January 7, 2014.

Objective

The primary objective of this research study was to explore and capture individual and small business taxpayer experiences with and expectations for ACA information that is provided via the IRS automated phone line. The results provided the IRS with insight into preferences, needs, and expectations in an effort to identify areas that needed to be remedied to provide a better overall experience for the taxpayer when seeking information. While these results are applicable to the ACA automated line, the IRS also developed best practices and general lessons learned that are applicable to other automated phone systems.

Research Method

Testing Locations

In order to collect data from a geographically diverse participant group, the IRS selected four cities to host testing sessions during August 2014. Researchers conducted a total of twelve testing sessions (three per city) in Washington, D.C.; Austin, Texas; Denver, Colorado; and Atlanta, Georgia (see Figure 1). Selected locations were limited to cities with IRS facilities or IRS-approved Federal facilities that could accommodate the following requirements:

- The facility had a room that could accommodate 10–15 participants for each 90-minute session.
- The room was available for at least one late afternoon session.
- Facilities could provide 10–12 outgoing phone lines with a minimum of 10 phones for each testing session.

FIGURE 1. Location of Testing Sessions



Participant Recruitment

The IRS used a contractor to recruit participants for each city drawing from a random sample of potential participants who met the recruitment criteria. For this study, all participants had to be 18 years of age or older and have:

1. received wage, pension, or annuity income in Tax Year 2013 for services performed as an employee of the U.S. Government or of any U.S. State or local government; and
2. filed a Tax Year 2013 tax return between cycles 1 and 13 in Processing Year 2014. This means the participants filed their returns by the established deadline for the processing year.

The IRS stratified the sampling pool based on gender, age, filing status, tax return preparation method, Business Operating Division (BOD) code, and Adjusted Gross Income (AGI). The BOD code indicated whether the taxpayer was to be placed into the small business testing session versus the individual testing sessions. Using the final sample, the contractor contacted taxpayers via phone in each city using a screener. The screener included questions to confirm the participant met all recruitment criteria and included questions about previous IRS service channel use and previous ACA-related knowledge and interactions. The contractor recruited a total of 119 taxpayers (80 individual participants and 39 small business participants) for the study.

Demographics

The participant pool was randomly selected to reflect diversity in demographic characteristics such as age, AGI, tax return preparation method, filing status and BOD code.

Participants were evenly spread among each age bracket ranging from 18 years old to 65 years and older (Table 1). Half of the participant population was represented in AGI brackets under \$50,000 (Table 2). Study participants' tax return preparation methods were evenly split between using a paid preparer (42 percent) and self-preparing an electronic return (49 percent) (Table 3). Half of the participant population had a "single" filing status (Table 4).

TABLE 1. Age of Study Participants

Age (in years)	Which category best describes your age?		
	Individual Participants	Small Business Participants	All Participants
18–25	10%	5%	8%
26–34	21%	12%	18%
35–44	19%	16%	18%
45–54	12%	16%	14%
55–64	22%	26%	23%
65 or older	16%	26%	19%

TABLE 2: Adjusted Gross Income of Study Participants

Adjusted Gross Income	Which category best describes your adjusted gross income?		
	Individual Participants	Small Business Participants	All Participants
\$0 to \$15,000	25%	26%	25%
\$15,001 to \$25,000	14%	7%	11%
\$25,001 to \$50,000	26%	23%	25%
\$50,001 to \$75,000	17%	9%	15%
\$75,001 to \$100,000	6%	7%	6%
Over \$100,000	12%	28%	18%

TABLE 3. Return Preparation Method of Study Participants

Preparation Method	Which category best describes your preparation method?		
	Individual Participants	Small Business Participants	All Participants
Paid Preparer	33%	58%	42%
IRS Sponsored Program (e.g., VITA)	1%	0%	1%
Self-Prepared, Paper	10%	7%	9%
Self-Prepared, Electronic	56%	35%	49%

TABLE 4. Filing Status of Study Participants

Filing Status	Which category best describes your filing status?		
	Individual Participants	Small Business Participants	All Participants
Head of Household	21%	14%	18%
Married Filing Joint	29%	28%	29%
Married Filing Separate	4%	0%	2%
Single	45%	58%	50%
Widow(er)	1%	0%	1%

Testing Sessions

Each testing session consisted of a three-part testing sequence that began with calling the main IRS phone line of 1-800-TAX-1040. Participants completed comprehension testing, evaluated their experience using a toll-free ratings sheet, and participated in a focus group discussion.

(A) Comprehension Testing

To facilitate the hands-on toll-free experience, participants were provided a series of fictional scenarios to use while calling the toll-free line and asked to answer comprehension questions aimed at measuring how well the participants understood the information provided on the automated phone line. Each set of comprehension questions was based on the specific scenario. The four scenarios provided to each participant mirrored issues that could be addressed using the ACA toll-free line (Table 5). Scenarios differed for individual taxpayer participants and small business taxpayer participants.

TABLE 5. Participant Comprehension Testing Scenarios

Individual	Small Business
<p>Scenario #1: "Only One Insured"</p> <ul style="list-style-type: none"> • Primary source of income for family • Job provides health insurance for only you • Need to determine the rules for health insurance coverage for your spouse and children • Need to determine if your whole family needs health insurance to avoid penalty 	<p>Scenario #1: "Healthcare Credit Qualifier"</p> <ul style="list-style-type: none"> • You own a small cookie company • Have obtained health coverage for your employees • You heard that some businesses might qualify for the health insurance tax credit • You need to determine if your business will qualify for the credit in 2015
<p>Scenario #2: "Penalty Avoidant"</p> <ul style="list-style-type: none"> • You do not have health insurance • You heard people can purchase insurance through the Marketplace to avoid penalty • If it's too late to avoid the penalty for next spring • You contact the IRS before going to the Health Insurance Marketplace 	<p>Scenario #2: "Insure 10 Employees"</p> <ul style="list-style-type: none"> • Opening a small boutique that will employ approximately 10 employees • Accountant informed you that you'll need to provide insurance for these employees • Accountant provided the IRS toll-free number and suggested you begin there

TABLE 5. Participant Comprehension Testing Scenarios—Continued

Individual	Small Business
<p>Scenario #3: "About Exemptions"</p> <ul style="list-style-type: none"> You do not have health insurance You don't make much money and heard that exemptions can be made on certain circumstances Need to know exemption options so you will not have to purchase coverage and avoid penalty You decide to contact the IRS to see if you qualify 	<p>Scenario #3: "Wondering about SHOP"</p> <ul style="list-style-type: none"> Own a small lawn care business Heard on the radio that all small business owners should be aware of something called SHOP The ad mentioned that SHOP can help business owners and something about taxes You decide to try contacting the IRS for more information
<p>Scenario #4: "Can't Afford"</p> <ul style="list-style-type: none"> You and your spouse cannot afford health insurance You want to avoid being penalized when it's time to file your 2014 tax return You contact IRS to get more information about potential options to avoid penalty 	<p>Scenario #4: "Non-profit Credit"</p> <ul style="list-style-type: none"> Owner of a small non-profit organization Provides health insurance as part of their employee benefit plan May qualify for a health insurance tax credit Decides to contact the IRS for more information about this credit and how to qualify for it

Participants were asked to record each numeric option pressed while attempting to collect the information needed to complete each scenario (Figure 2).

FIGURE 2. Participant Recording Sheet

OMB# 1545-1349

Participant number: _____

Scenario number: _____

Call Route:

Pressed: __

➔

Pressed: __

Were you disconnected? (circle one) **Yes** **No**

Did you reach an assistor? (circle one) **Yes** **No**

Notes:

Participants were asked to answer a series of comprehension questions after completing each scenario. The hands-on experience was self-guided for 35 minutes. All participants were provided the same toll-free number (1-800-TAX-1040), navigated themselves through the provided ACA information, and were allowed to attempt each scenario as many times as they desired in order to fully complete the comprehension questions.

(B) Toll-Free Ratings Sheet

After completing the comprehension testing portion of the session, participants were given 10 minutes to complete a questionnaire evaluating their experience using the toll-free automated line. Participants were asked to rate various aspects of their experience using the ACA toll-free automated phone line. Using a 4-point rating scale, participants responded to the items shown in Figure 3.

FIGURE 3. Participant Toll-Free Ratings Sheet²

Toll-Free Ratings Sheet		Participant Number: _____			
1. Please rate your satisfaction with your overall phone experience.	Not at all Satisfied	Slightly Satisfied	Very Satisfied	Completely Satisfied	<input type="radio"/>
2. How easy or difficult was it for you to use the phone system to complete the tasks?	Very Difficult	Difficult	Easy	Very Easy	<input type="radio"/>
3. Please rate your experience with each of the following aspects of the phone system:	Poor	Fair	Good	Excellent	
Number of menu options	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Speed of the recorded message	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ability to understand the recorded message	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
4. How useful was the information you received from the phone system?	Not at all Useful	Not very Useful	Moderately Useful	Very Useful	<input type="radio"/>
5. Rate the quality of the information you received from the phone system.	Poor	Fair	Good	Excellent	<input type="radio"/>
6. Please rate your satisfaction with the amount of time it took to complete each task.	Not at all Satisfied	Slightly Satisfied	Very Satisfied	Completely Satisfied	<input type="radio"/>
7. Based on your experience, would you be willing to use this service in the future?					
<input type="radio"/> Yes <input type="radio"/> No					

(C) Focus Group Discussion

During the remaining 35 minutes of each testing session, participants were asked to share their overall user experience and recommendations during a moderated focus group discussion. The qualitative data collected through the focus group portion of the testing session was used to understand the taxpayer experience of using the ACA toll-free line and to obtain common themes among participants. The discussion was moderated by an IRS Research Analyst.

(D) Collection and Analysis of Data

Participant responses from the comprehension testing portion of the testing sessions were totaled across scenarios. The responses were totaled separately for individual and small business participants. The toll-free rating worksheet responses for both individual and small business participants were coded and validated. The focus groups were recording and transcribed, and the data were analyzed for similar thoughts and ideas to be categorized into themes. The data were evaluated collectively for all cities combined, but were analyzed separately for individual and small business participants. Across the 4 cities, 119 participants completed sessions.

² Toll-Free Ratings Sheet scale was developed by Internal Revenue Service (IRS) Wage & Investment Research and Analysis (WIRA) Research Group 4, June 9, 2014.

(E) Limitations

When reviewing this research, it is important to take into consideration that participants may respond and behave to hypothetical scenarios and situations significantly differently than they would when faced with a real-life, personal situation. This disclaimer may also be applicable to their established expectations for provided services, actual usage methods of services, understanding the context of the information provided to them in the scenarios as well as countless other factors. When conducting research such as this, it is important to control for as many of those factors as possible. Going forward, several lessons learned can be applied to future research.

The participants responded to a questionnaire during recruitment about their existing knowledge and previous experiences regarding ACA-related information. Though those responses were captured, they were not used in any way to divide the testing participants into experienced and novice categories. In future studies, it would be interesting to measure the differences in overall testing experiences across the varying levels of previous knowledge and ACA-related interactions. All individual participants were provided the same four testing scenarios, and similarly, all small business participants were provided the same four testing scenarios, based on the IRS automated line content. Not all participants connected to the content of the provided scenarios, which resulted in complications and confusion during testing. Often, participants did not complete all four scenarios due to confusion and the limited time frames. In future research, it would be beneficial to alter the content of the scenarios to be more relatable based on current and relevant issues as well as adjust the way the scenarios themselves, as well as the number of scenarios, are administered.

As shown in Figure 2, participants were asked to record the menu items they selected as they were going through the prompts on the automated line, using the Participant Recording Sheet. Because of the size of the groups (approximately 10 participants in each session), it was difficult for the research team to monitor the participants to ensure that this task was being completed correctly. Once the team began to analyze the data, it became apparent that overall the participants had difficulty completing this portion. Because of the significant inconsistencies with the responses for this portion, the data could not be used for any results outside of noting that the task was difficult. The research team was hoping to record the pathways that the participants followed. In future research, steps should be taken to ensure that this task can be completed successfully, possibly using smaller groups or more session assistants.

Results

(A) Comprehension Testing

The responses received by individual participants are shown below in Tables 6 and 7. Not all taxpayers completed all scenarios, so the base number of responses to each question is unique. With each scenario, taxpayers were asked the following questions:

- Were you able to resolve your issue in the time allotted?
- Did you want to speak to an assistor?
- Were you disconnected?
- Did you reach an assistor?

TABLE 6. Individual Experience Attempts

Individual Experience	Scenario I-1	Scenario I-2	Scenario I-3	Scenario I-4	All Individual Attempts*
Disconnected	26%	16%	7%	20%	18%
Reached an assistor	6%	6%	0%	7%	5%

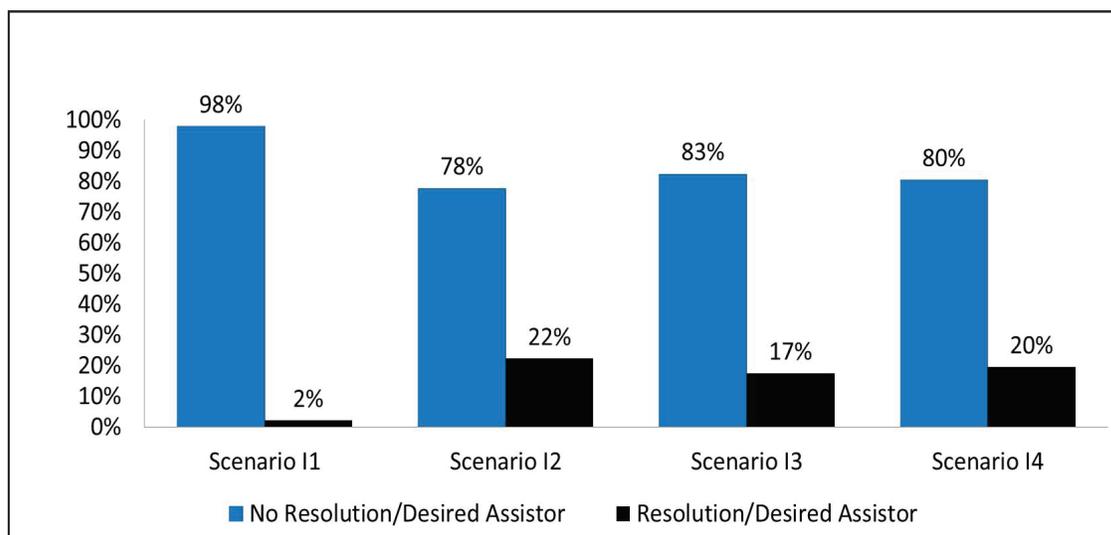
* Since participants completed the scenarios at their own pace, the base number of responses varied across scenarios.

TABLE 7. Individual Comprehension Results

Individual Comprehension	Scenario I-1	Scenario I-2	Scenario I-3	Scenario I-4	All Individual Responses*
Resolved issue	12%	38%	26%	19%	24%
No issue resolution	88%	62%	72%	81%	76%
Desired an assistor	72%	67%	53%	67%	65%

* Since participants completed the scenarios at their own discretion and pace, the base number of responses varied across scenarios..

Approximately 18 percent of all individual attempts to complete the scenarios resulted in a disconnected call. Across the four scenarios, 12 percent to 38 percent of individual participants reported successfully resolving their issue, while 62 percent to 88 percent of individual participants reported not being able to resolve their issue. Overall, the majority (65 percent) of responses from individual participants indicated a desire to speak to an assistor while trying to work through the testing scenarios. However, only a small proportion (5 percent) of all individual attempts to complete the scenarios resulted in participants reaching an assistor. Of all individual participants who desired to speak to an assistor, those who reported being unable to resolve their issue were more likely to indicate a desire to speak to an assistor than those individual participants who resolved their issue, as seen in Figure 4.

FIGURE 4. Individual Participants/Assistor Desired

As previously mentioned, both individual and small business participants responded to a series of questions. The responses from the small business participants are shown below in Tables 8 and 9.

TABLE 8. Small Business Experience Attempts

Small Business Experience	Scenario SB-1	Scenario SB-2	Scenario SB-3	Scenario SB-4	All Small Business Attempts*
Disconnected	19%	14%	13%	13%	15%
Reached an assistor	34%	26%	13%	67%	30%

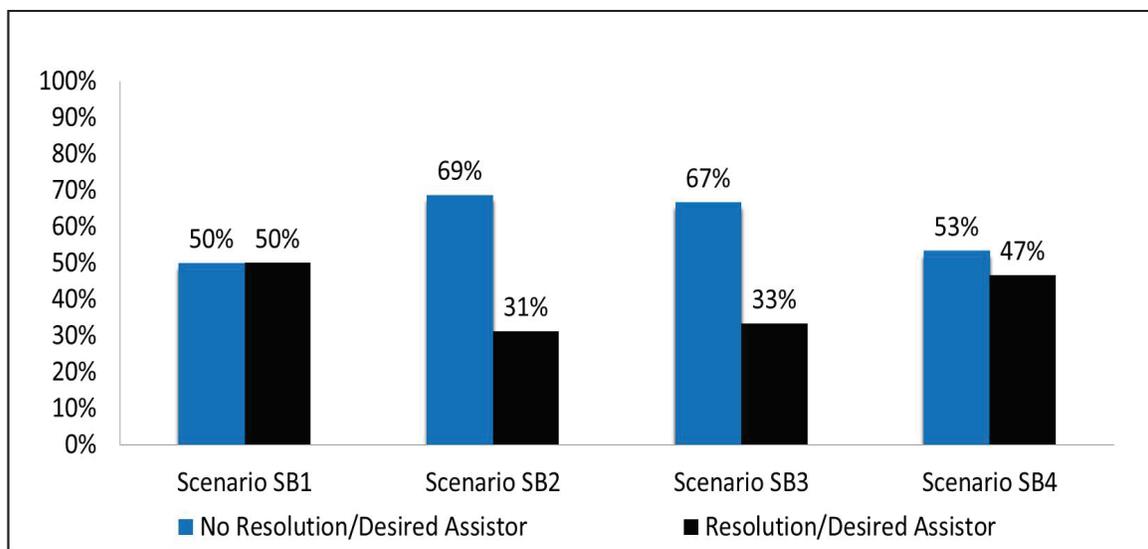
* Since participants completed the scenarios at their own discretion and pace, the base number of responses varied across scenarios.

TABLE 9. Small Business Comprehension Responses

Small Business Comprehension	Scenario SB-1	Scenario SB-2	Scenario SB-3	Scenario SB-4	All Small Business Responses*
Resolved issue	53%	47%	54%	61%	54%
No issue resolution	47%	53%	46%	39%	46%
Desired an assistor	49%	43%	37%	40%	42%

* Since participants completed the scenarios at their own discretion and pace, the base number of responses varied across scenarios.

Approximately 15 percent of all small business participant attempts to complete the scenarios resulted in a disconnected call. Across the four scenarios, 47 percent to 61 percent of small business participants reported successfully resolving their issue, while 39 percent to 53 percent of small business participants reported not being able to resolve their issue. Overall, 42 percent of the responses from small business participants indicated a desire to speak to an assistor while trying to work through the testing scenarios. A total of 30 percent of all small business participant attempts to complete the scenarios resulted in the taxpayer reaching an assistor. Of all small business participants who desired to speak to an assistor, those who reported being unable to resolve their issue were more likely to indicate a desire to speak to an assistor than those small business participants who resolved their issue, as seen in Figure 5.

FIGURE 5. Small Business Participants-Assistor Desired

Small business participants were more likely to report being able to resolve their issue than individual participants. Additionally, fewer small business than individual participants indicated a desire to speak to an assistor while trying to work through the testing scenarios (42 percent vs. 65 percent), and among those who indicated such a desire there was greater variability in issue resolution across scenarios among small business participants compared with similar individual participants. However, the desire to speak to an assistor was well represented among both individual and small business participants, with that desire generally even stronger for those participants reporting not having reached issue resolution.

Overall, both individual and small business participants experienced difficulty navigating the toll-free line, and often were not able to reach the intended or expected information. Please reference Appendices A and B for a detailed breakdown of responses provided to the comprehension questions for each scenario.

(B) Toll-Free Ratings Sheet

A total of 119 taxpayers completed the toll-free ratings sheet: 80 individual participants and 39 small business participants. Individual participants reported not being satisfied with their overall automated phone experience (57 percent not at all satisfied, 43 percent slightly satisfied), and reported varying levels of difficulty with using the automated phone service to complete the tasks (27 percent very difficult, 33 percent difficult). Small business participants were more satisfied with the automated phone service (56 percent slightly satisfied, 26 percent very satisfied), and reported that it was easy to use the tool to complete the tasks (61 percent easy, 16 percent very easy). According to the North American Consumer Technographics Customer Life Cycle Survey 2,³ which was administered in 2014, some 48 percent of U.S. online adults indicate using voice self-service in the past 12 months, and among those who used that service, 56 percent indicate being satisfied with their experience using this method. Similarly, the participants in the current study reported a 56-percent level of satisfaction (including the responses of slightly satisfied, very satisfied and completely satisfied, as seen in Table 10).

Overall, approximately 33 percent of all participants were willing to use the automated phone service again, while 67 percent of all participants reported being unwilling to use the automated phone service in the future (Table 11).

TABLE 10. Participant Toll-Free Ratings Sheet

Overall Satisfaction	Please rate your satisfaction with your overall phone experience.		
	Individual Participants	Small Business Participants	All Participants
Not at all satisfied	57%	15%	44%
Slightly satisfied	43%	56%	47%
Very satisfied	0%	26%	8%
Completely satisfied	0%	3%	1%

TABLE 11. Participant Toll-Free Ratings Sheet

Future Willingness To Use	Would you be willing to use this service again?		
	Individual Participants	Small Business Participants	All Participants
Yes	12%	76%	33%
No	89%	24%	67%

(C) Focus Group Discussion

The qualitative testing session data were collected by the IRS Research Team through focus groups to better understand the taxpayer experience using the ACA toll-free line and to obtain common themes. The quotes provided below are representative of the overall responses that were received from all locations. Each supporting quote has been notated with a location and time to identify the session in which the specific comment was heard.

Participants do not understand the IRS's role in implementing ACA in comparison to other agencies.

- “The IRS are the people who are handling the Affordable Care Act? I didn't know that.”⁴
- “I think if they could, it would be better if they separated. If you wanted to find out about healthcare that you wouldn't be working, doing anything, with the IRS at all unless you had a specific question about your taxes, having to do with healthcare.”⁵

³ North American Consumer Technographics Customer Life Cycle Survey 2, 2014. Survey administered by Forrester Research, Inc.

⁴ Austin, TX, Focus Group, 8/20/2014, 2:00PM

⁵ Denver, CO, Focus Group, 8/22/2014, 11:00AM

When calling the toll-free line, participants would prefer to resolve their issue within that service channel. Many participants expressed a desire to speak with a live assistor, rather than the automated line, but understood that there are constraints.

- “If I’m calling the phone number, it’s because I want information from the phone number. If I was going to go to the website, I would’ve started with the website.”⁶
- “To have every question answered by a real, live person. We can’t afford that. It’s a nice idea, but I call up and want to know what the office hours are. Should somebody be there answering the telephone and give me a human voice to tell me what the office hours are? No, there can just be a machine that tells me that.”⁷

Participants’ expectations impact satisfaction. Altering front matter (initial scripts taxpayers hear) to help establish educated and realistic expectations for the taxpayers’ experience can help alleviate confusion and frustration.

- “There’s a lot of people that just don’t know how to do that (navigate the website). They’re hoping, when you call, you either get a live person or get those answers through the prompts.”⁸
- “I expected that I could push buttons and get a recorded answer to the question. That didn’t happen. Then I thought I’d probably be able to find a person. That didn’t happen, and then I thought, ‘Am I doing this right?’”⁹

The participants have a genuine desire to be compliant, but require more detailed information. Many taxpayers expressed that they need more information about how they are impacted and next steps.

- “You need more meat. You need more meat to what’s going on.”¹⁰
- “You’re calling to gain information. Some of what she was telling you could answer your question. But if it doesn’t answer your question, you kind of zone out, trying to get to-- how do I get my question answered?”¹¹

Making other information channels available would be appreciated.

- “I think maybe even better than this approach would be a live chat. You know how you do that, when you’ve got computer problems, and you can type in something? You usually get a response a lot quicker.”¹²
- “Call back, that’s a good idea. Automatically call you back when they’re ready.”¹³

Changes to certain logistical elements of the phone tool will create a smoother user experience.

- “When it comes to giving out that detailed information like that, just slow it down.”¹⁴
- “I found the computer voice distracting, as opposed to a real recorded person’s voice.”¹⁵
- “And there was no prompt to repeat something. It was a long sentence. If you forgot the last part or you weren’t paying attention to the last part, all you would have to do is repeat again, but you couldn’t because you’d have to go back through the steps.”¹⁶

⁶ Austin, TX, Focus Group, 8/20/2014, 11:00AM

⁷ Austin, TX, Focus Group, 8/20/2014, 2:00PM

⁸ Denver, CO, Focus Group, 8/22/2014, 11:00AM

⁹ Austin, TX, Focus Group, 8/19/2014, 3:00PM

¹⁰ Atlanta, GA, Focus Group, 8/29/2014, 9:30AM

¹¹ Washington, D.C., Focus Group, 8/14/2014, 11:00AM

¹² Denver, CO, Focus Group, 8/22/2014, 2:00PM

¹³ Atlanta, GA, Focus Group, 8/29/2014, 9:30AM

¹⁴ Washington, D.C., Focus Group, 8/13/2014, 3:00PM

¹⁵ Denver, CO, Focus Group, 8/21/2014, 3:00PM

¹⁶ Atlanta, GA, Focus Group, 8/29/2014, 9:30AM

Active Changes as a Result of the Study

Members from the IRS Research Team collaborated with representatives from several business units within the IRS to improve the existing ACA-related call scripts and routing based on experience and insights gained during the testing sessions. The revised scripts set taxpayer expectations up front related to live assistance, helping ease and/or avoid the frustration of taxpayers who can unnecessarily find themselves caught in an endless loop of automated messages while trying to reach a live assistor. Following implementation of the revised scripts, analysis of telephone data from December 2014 through February 2015 reveals that the percentage of calls dropped before the completion of the up-front message dramatically increased from 8 percent to 26 percent. With the IRS' toll-free telephone environment programmed to block calls when the queue reaches a defined threshold (referred to as a courtesy disconnect; taxpayers receive a recorded announcement to call back at a later time), the increase in the percentage of calls dropped during the up-front messaging allows the IRS to provide service to more taxpayers. This success of evaluating the taxpayer experience prior to releasing new phone content has facilitated long-term opportunities for the team to apply this research approach to revise other IRS phone scripts.

Conclusions

Faced with these budgetary constraints and high call demand, the IRS has increasingly had to rely on automated options to handle taxpayer needs including pre-recorded telephone scripts related to general tax information, changes in tax laws, etc. This research provides insight for improvements to automated telephone messages and associated content with the goal of striking the optimal balance of providing the necessary and relevant amount of information through automation, which will supply taxpayers with the information needed to answer their questions and/or resolve their issues without needing to wait in queue to speak with a live assistor. Furthermore, the findings from this research helped inform changes to the content, organization, and delivery of information on the ACA automated line. As mentioned in the limitations, if the research is ever duplicated there are several areas where improvements can be made to better capture a genuine taxpayer experience. However, for a first attempt, this research was successful.

Acknowledgments

Directors and Management

Toni Cross, Director of WIRA

Mia Sylve, Acting Director of WIRA

Elizabeth L. Blair, Senior Operations Advisor to the Director of WIRA

David C. Cico, Chief of WIRA Research Group 4

Javier A. Framinan, Chief of WIRA Research Group 2

Sasha Lanes, Chief of WIRA Research Group 3

Core Project Team

Ariel S. Wooten, Lead Social Scientist

Marisa E. McDaniels, Lead Operations Research Analyst

Julie Thompson-Evans, Contracting Officer Representative (COR)

Patti Davis-Smith, Social Scientist

Maria C. Wang, Operations Research Analyst

Robert P. Thomas, Social Scientist

David C. Cico, Chief of WIRA Research Group 4

Contributing Researchers

Courtney Rasey, Technical Lead Social Scientist Research Group 4

Howard Rasey, Technical Lead Social Scientist Research Group 3

Saima Mehmood, Social Scientist Research Group 2

Mackenzie Wiley, Social Scientist Research Group 2

Lisa Elliott, Social Scientist Research Group 3

Contractor Assistance

Fors Marsh Group (FMG), Arlington, VA

Appendix A

Individual Participant Testing Scenarios and Comprehension Findings

OMB #:1545-1349

Participant Number: _____

Scenario I1

You're the primary source of income for your family and your job provides health insurance for only you, free of charge. You need to find out what the rules are regarding health insurance coverage for your spouse and children; in other words, does your whole family need health insurance coverage to avoid the penalty for not having coverage? Your spouse suggested that you contact the IRS for more information. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

Will your entire family need coverage to avoid the penalty?	Number of similar responses
Yes	6
Go to the website	4
I don't know	26

What if you need additional information?	Number of similar responses
IRS.gov	48
Healthcare.gov	24

OMB #:1545-1349

Participant Number: _____

Scenario I2

You're an individual who does not currently have health insurance. You heard that people can purchase insurance through the Health Insurance Marketplace to avoid the penalty for not having coverage. You want to start the process today but do not know if it is too late to avoid the penalty for next spring. You decide to contact the IRS before going to the Health Insurance Marketplace. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

Is it too late to purchase insurance to avoid the 2014 penalty?	Number of similar responses
Yes	14
No	4
I don't know	12

What if you need additional information?	Number of similar responses
IRS.gov	18
Healthcare.gov	10

OMB #:1545-1349

Participant Number: _____

Scenario I3

You're an individual who does not currently have health insurance coverage, but heard from some co-workers that there are exemptions that can be made based on certain circumstances. Because you are not making much money in your current job, you hope that an exemption would allow you to not have to purchase coverage and also exempt you from the penalty for not having coverage. You decide to contact the IRS to see if you can qualify for an exemption. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

Will you be able to file your 2014 tax return if you do not have insurance?	Number of similar responses
Yes	10
No	4
I don't know	10

What if you need additional information?	Number of similar responses
IRS.gov	4
Healthcare.gov	22

OMB #:1545-1349

Participant Number: _____

Scenario I4

You and your spouse cannot afford health insurance and believe that you will be penalized when it's time to file your 2014 tax return. Your neighbor suggested that you contact the IRS to get more information about potential options to avoid the penalty. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

What can you do to avoid the penalty?	Number of similar responses
Go to the website	6
Get insurance	6
Depends on if we qualify for exemptions	4
I don't know	18

What if you need additional information?	Number of similar responses
IRS.gov	6
Healthcare.gov	28

Appendix B

Small Business Participant Testing Scenarios and Comprehension Findings

OMB #:1545-1349

Participant Number: _____

Scenario SB1

You own a small cookie company and have obtained health coverage for your 3 employees. You heard from a friend (another bakery owner) that some businesses might qualify for the health insurance tax credit. You need to determine if your business will qualify for the credit in 2015 and decide to contact the IRS for more information. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

Do you need to requalify your business for 2015?	Number of similar responses
Yes	15
Yes, but I want to see what the specific requirements are	1
No	6
I don't know	13
Was diverted to the website 6 times with no response provided	1

What if you need additional information?	Number of similar responses
IRS.gov/aca	22
Call back	4
Try to find SHOP website	2
Healthcare.gov	1
I don't know, try to seek information from another source	2

OMB #:1545-1349

Participant Number: _____

Scenario SB2

You're in the process of opening a small boutique that will employ approximately 10 employees. Your accountant told you that you'll need to provide insurance for these employees but you don't know where to begin. Your accountant provided the IRS toll-free number and suggested you begin there. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

What action did the toll-free line tell you to do in order to insure employees?	Number of similar responses
IRS.gov/aca	12
Healthcare.gov and search for minimum essential coverage	9
Talk to an assistor	2
I don't know	5

What if you need additional information?	Number of similar responses
IRS.gov/aca	16
Healthcare.gov, search for minimum essential coverage	6
Call SHOP	3
I don't know	6

OMB #:1545-1349

Participant Number: _____

Scenario SB3

You own a small lawn care business and heard on the radio that all small business owners should be aware of something called SHOP. The ad mentioned that SHOP can help business owners and something about taxes. Because you were driving, you did not hear the telephone number or website the commercial provided. You decide to try contacting the IRS for more information. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

What does SHOP stand for?	Number of similar responses
Small Business Option Program	26
I don't know	5
Can't remember, wasn't told multiple times	3

What if you need additional information?	Number of similar responses
IRS.gov/aca	17
Healthcare.gov	10
Call SHOP	3
I don't know	6

OMB #:1545-1349

Participant Number: _____

Scenario SB4

You own a small non-profit organization that provides meals for the elderly and you currently provide health insurance as part of your employee benefit plan. One of your employees told you that you may qualify for a health insurance tax credit because of the organization's tax-exempt status. After chatting with your employee about this information, you decided to contact the IRS for more information about this credit and how to qualify for it. Your task today is to try to resolve your issue using the following toll-free number: 1-800-829-1040.

Where must you purchase for your employees to qualify for the health insurance tax credit?	Number of similar responses
Small Business Option Program	21
IRS.gov/aca	3
Through the employment tax department	1
I don't know	4

What if you need additional information?	Number of similar responses
IRS.gov/aca	25
Healthcare.gov	4

4



Helping Taxpayers Get It Right

Gleason ♦ Tong

Maag ♦ Edelstein ♦ Hanson ♦ Minton ♦ Pergamit ♦ Ratcliffe

Masken

Nontaxable Combat Pay Election and the Earned Income Tax Credit

Suzanne Gleason and Patricia K. Tong,
Office of Tax Analysis, U.S. Department of the Treasury¹

1. Introduction

It is no secret that the U.S. Federal income tax system has become increasingly complex as policy makers use it to provide income support to targeted populations and incentives to support certain behaviors. In light of this continuing trend, the tax policy community should further its understanding of whether taxpayers are benefiting from these programs or are simply too confused to figure out how to take advantage of the tax benefits being made available. In addition, policymakers need to determine whether taxpayers can navigate the tax code on their own or if a paid preparer is required for tax minimization. In this paper, we shed light on these questions by examining the use of the nontaxable combat pay election (NCPE) by military personnel eligible to claim the Earned Income Tax Credit (EITC), a refundable tax credit targeted to lower income working families.

Combat pay includes all military pay—including wages earned as well as any reenlistment or other bonuses, etc.—received during a month in which a service member was stationed in a combat zone for at least one day. Combat pay is exempt from income taxes, thus providing military personnel with additional compensation for serving in dangerous areas.² Prior to 2005, combat pay was not counted as earned income for the purposes of calculating the EITC. In 2004, the U.S. General Accounting Office (GAO) issued a report demonstrating that the exclusion of combat pay from the EITC calculation reduced the EITC for some of the lowest income members of the armed forces while increasing EITC eligibility among others, generally with higher total incomes. As a result, the Working Families Tax Relief Act of 2004 gave taxpayers the option (election) to include combat pay in EITC earned income. The **Heroes Earnings Assistance and Relief Tax Act of 2008** made this change permanent. In this study, we analyze the use of this election. There is evidence that taxpayers may not fully understand the relationship between their earnings and the amount of EITC that they could claim (Chetty and Saez, 2013). Due to the complex structure of the EITC itself, taxpayers may not understand when to include or exclude combat pay in order to minimize their income taxes.

In this paper, we combine administrative tax data and Department of Defense (DOD) personnel data to determine the extent to which military service personnel make the optimal election (i.e., correctly use combat pay to maximize their EITC). We examine the demographic and financial characteristics of these taxpayers for each category of optimization behavior. In addition, we estimate the extent to which the exclusion of combat pay increases EITC eligibility among personnel serving in combat zones. Finally, we estimate the aggregate cost associated with the NCPE and the cost of increasing eligibility to personnel who would not otherwise qualify for the EITC.

We find that offering taxpayers a choice about how to treat their combat pay for the purposes of earned income in the EITC calculation results in some taxpayers failing to make the optimal decision. We also show that taxpayers are less likely to make the tax-minimizing choice if they are low income, if they prepare their own tax returns, or when the optimal choice is more complicated. Although the NCPE has a small effect on the total cost of the combat pay exclusion, it does affect the distribution of those benefits among similarly positioned military personnel.

¹ Views and opinions expressed are those of the authors and do not necessarily represent official Treasury positions or policy. This work was supported by the U.S. Department of Defense. The authors are grateful to the U.S. Department of Defense and the Internal Revenue Service for providing data for this study. Many thanks to seminar participants at the 2011 Association for Public Policy and Management Fall Research Conference, 2011 National Tax Association Annual Conference on Taxation, 2012 American Economic Association Annual Meeting, and 2015 IRS-TPC Research Conference.

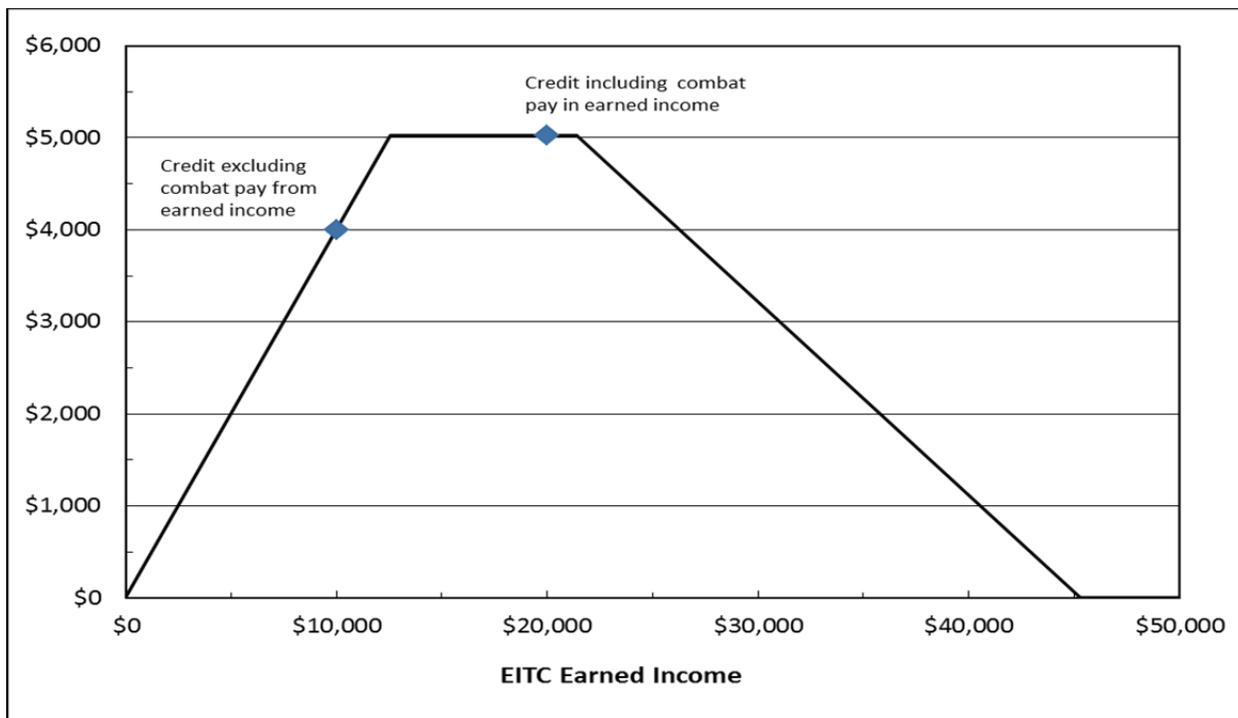
² Combat pay is exempt from income taxes, but is subject to payroll taxes.

2. Background

The EITC is a refundable tax credit that provides work incentives to lower income households. In Tax Year 2009, about 27 million families claimed the EITC, representing almost 20 percent of all tax returns filed (Statistics of Income, 2011). The EITC schedule has a trapezoidal shape in which the credit phases in with earnings, plateaus, and then phases out with the greater of earnings and adjusted gross income (AGI). The EITC schedule depends on filing status and number of EITC qualifying dependents. Single filers with no dependents have the least generous EITC schedule and were eligible for a maximum credit of \$457 in 2009. Married couples filing jointly with three or more EITC qualifying children have the most generous EITC schedule, being eligible for a maximum credit of \$5,657 in 2009.

Figure 1 provides an example of how the NCPE can increase the amount of EITC claimed by a taxpayer. If a taxpayer includes nontaxable combat pay in EITC earned income, then he is required to include the entire amount of it; he can't include just a portion of his combat pay to achieve an even greater credit amount. The figure presents the 2009 EITC schedule for a married couple filing jointly with two children who have \$10,000 in taxable wages and \$10,000 in nontaxable combat pay. By default, nontaxable combat pay is excluded from EITC earned income. Thus, if the family does not use the NCPE, then they would claim an EITC of \$4,000. However, if the family were to include their nontaxable combat pay, then their EITC earned income increases to \$20,000 and their EITC would increase by over \$1,000 from \$4,000 to \$5,028. This example demonstrates the potentially significant reductions in EITC when nontaxable combat pay is excluded from EITC earned income.

FIGURE 1. 2009 Earned Income Tax Credit for a Married Couple Filing Jointly With Two Children, \$10,000 in Taxable Wages, and \$10,000 in Combat Pay



3. Data

For this study, we constructed an individual-level dataset containing the population of EITC-eligible military personnel with nontaxable combat pay by combining military personnel data with tax return data for Tax Years 2005-2009. Military personnel data come from the U.S. Department of Defense (DOD) and include characteristics such as pay grade, service branch (Army, Navy, Marines, or Air Force), and component (active, reserve, or National Guard). Base salary, also known as the pay grade, for military personnel is determined by service, rank and tenure. Individual tax return and Form W-2 wage data from the Internal Revenue Service (IRS) Compliance Data Warehouse (CDW) were matched to the DOD personnel data. The tax return data consist of financial items reported on the IRS Form 1040, as well as filing characteristics such as filing status and the use of a paid preparer. Because tax return data are reported by

tax unit (which sometimes includes a spouse and children), variables such as AGI are aggregated across individuals within a tax unit. In contrast, W-2s are reported at the individual level. W-2s for each person in the DOD personnel data were extracted and summed across employers to calculate total W-2 wages. As of 2005, nontaxable combat pay is reported by the military in Box 12 of Form W-2 using Code Q.

The sample is restricted to personnel with nontaxable combat pay reported on their W-2s who are eligible to claim the EITC.³ Personnel are considered eligible if they satisfy all the EITC requirements listed in the IRS Form 1040 instructions, or if they claim the EITC on their tax return.⁴ The sample consists of 987,454 individuals, representing over 30 percent of the personnel with nontaxable combat pay over the 5-year period. Each individual represents one tax unit. For filers who claim the EITC, we take the number of EITC qualifying dependents as given. For filers who do not claim the EITC, we assume that the number of EITC qualifying dependents equals the number of exemptions claimed for children living at home. The sample includes personnel who do not file a Federal tax return, but appear to be eligible for the EITC based on their W-2 wages. For nonfilers, a simplified tax return was constructed based on their W-2 data and DOD demographic characteristics, where W-2 wages were assumed to be the only source of taxable earned income and the number of dependents under age 17 was assumed to be the number of qualifying EITC children. Nonfilers represent 5 percent of the sample and are assumed to be nonoptimizers. Spouses who both report nontaxable combat pay are counted as two tax unit observations in the data. These dual-military couples make up less than 2 percent of the EITC-eligible military personnel population with nontaxable combat pay.

4. Analysis

4.1 Nontaxable Combat Pay Election

To determine if a taxpayer uses the NCPE in such a way as to maximize EITC, we calculated the EITC both including and excluding nontaxable combat pay in EITC earned income.^{5,6} These amounts were compared to the EITC calculated with the earned income reported on the IRS Form 1040. The taxpayer was considered an optimizer when the EITC calculated using his IRS Form 1040 reported earned income equaled the largest credit. Individuals who do not claim the EITC, including nonfilers, are classified as nonoptimizers. These calculations reveal that only 5.6 percent of EITC-eligible combat personnel take advantage of the NCPE. Given that nearly 75 percent have taxable earned income higher than the first EITC kink and the average amount of nontaxable combat pay is large (\$22,519), it is not surprising that such a small percentage of EITC-eligible combat personnel elect to include combat pay in their EITC earned income.

Table 1 provides characteristics for those who do not optimize the EITC in column 1 and for those who do optimize the EITC in column 2. More than 80 percent of those combat personnel eligible for the EITC maximize their credit. The distributions by service and pay grade are roughly the same across the two groups with the sample of optimizers containing fewer National Guard and reserve personnel. On average, nonoptimizing personnel are slightly older, more likely to file as single or head of household as opposed to married filing jointly, and have lower AGI than optimizing personnel.

Most nonoptimizers (62 percent) are filers who do not claim the EITC. The remaining nonoptimizing personnel are either nonfilers (31 percent), claim the EITC and use the NCPE (4 percent), or claim the EITC and do not use the NCPE (3 percent). By construction, all optimizers claim the EITC. Among optimizers, 6 percent use the NCPE. A higher percentage of optimizing personnel have at least one EITC qualifying dependent compared to nonoptimizing personnel (82 percent versus 48 percent). This difference in optimization rates is consistent with historical participation rates by number of children, which show that taxpayers eligible for the childless EITC have lower take-up rates than EITC-eligible taxpayers with children.⁷ Average nontaxable combat pay among nonoptimizing personnel is \$23,635,

³ Individuals with missing DOD personnel data were dropped from the sample. In addition, individuals who claim the EITC and have zero reported EITC earned income were excluded. These individuals represent less than 1 percent of the sample.

⁴ Personnel who do not originally claim the EITC, were deemed eligible if adjusted gross income and investment income were below means-tested thresholds and age and filing status requirements were satisfied. Furthermore, personnel not originally claiming the EITC were considered eligible only if they could claim a positive EITC with either the inclusion and/or exclusion of nontaxable combat pay in EITC earned income. The authors acknowledge that these requirements do not perfectly characterize the sample of EITC-eligible personnel due to unobservable idiosyncratic circumstances.

⁵ The continuous EITC formula was used to calculate the amount of EITC originally claimed and the counterfactual EITC. Taxpayers may report a different amount of EITC claimed because the IRS Form 1040 instructions refer taxpayers to a table to look up the amount of EITC.

⁶ For dual-military couples, an analogous methodology was applied to calculate the EITC under four mutually exclusive scenarios for each member of the couple: (1) exclude all combat pay; (2) include own combat pay; (3) include spouse's combat pay; and (4) include the sum of own combat pay and spouse's combat pay.

⁷ According to Plueger (2009), the take-up rate of the childless EITC was 56 percent, which is substantially lower than the take-up rates of those with 1 child (74 percent) and 2 or more children (86 percent) in Tax Year 2005. Similarly, a study by the U.S. General Accounting Office (2001) shows that the 1999 EITC take-up rate among childless taxpayers was significantly lower than the take-up rate among taxpayers with children.

which is higher than the average for optimizing personnel (\$21,877). Nonoptimizers are also less likely to use a paid preparer. Twenty-eight percent of nonoptimizers use a paid preparer, which includes people who incorrectly apply the NCPE as well as personnel who do not claim the EITC.

Table 2 depicts EITC optimization rates for subpopulations of personnel by optimal strategy and paid tax preparer use. Overall, 6.5 percent of the population should elect to include their combat pay to maximize their EITC. Even though the proportion of personnel using paid tax preparers is higher among personnel who should include their combat pay (57 percent versus 54 percent), their optimization rates are lower than personnel who should exclude their combat pay (71 percent versus 83 percent). Higher observed optimization rates among those who should exclude their combat pay could be partially explained by the design of the policy, which defaults to excluding nontaxable combat pay from the EITC calculation.

TABLE 1. Sample Summary Statistics by Earned Income Tax Credit (EITC) Optimization, Tax Years 2005–2009

Characteristics	(1) Do Not Optimize the EITC		(2) Optimize the EITC	
	Share or Mean	Standard Deviation	Share or Mean	Standard Deviation
Distribution Across Military Service Branches				
Army	72%		65%	
Air Force	09%		13%	
Marines	08%		10%	
Navy	11%		13%	
Total	100%		100%	
Distribution Across Military Component				
National Guard	21%		13%	
Reserves	14%		09%	
Active	66%		78%	
Total	100%		100%	
Distribution Across Pay Grades^a				
Enlisted: Pay Grade ME01-ME03	9%		7%	
Enlisted: Pay Grade ME04	31%		27%	
Enlisted: Pay Grade ME05	27%		29%	
Enlisted: Pay Grade ME06	15%		19%	
Enlisted: Pay Grade ME07-ME09	8%		10%	
Officer	10%		8%	
Total	100%		100%	
Other Characteristics				
Age	31.05	(7.19)	30.16	(6.51)
Single	29%		16%	
Married Filing Jointly	59%		74%	
Head of Household	12%		10%	
Any EITC Qualifying Dependents	48%		82%	
Earned Income Tax Credit	\$626	(1,249)	\$2,315	(1,561)
Adjusted Gross Income	\$12,127	(14,047)	\$17,522	(10,756)
Nontaxable Combat Pay	\$23,635	(16,004)	\$21,887	(14,970)
Paid Preparer	28%		60%	
Claim EITC	7%		100%	
Nonfiler	31%		0%	
Use Nontaxable Combat Pay Election	4%		6%	
Observations	172,808		814,559	
Share	17.5%		82.5%	

NOTE: Monetary variables are in 2009 dollars. Totals may not add to 100% due to rounding.

^a See footnote to Table 4 for pay grade salaries.

TABLE 2. Optimization Rates by Optimal Strategy and Use of a Paid Tax Preparer, Tax Years 2005–2009

Item	All	Optimal Strategy	
		Include Nontaxable Combat Pay	Exclude Nontaxable Combat Pay
Optimization Rate	82%	71%	83%
Average EITC Forgone Among Nonoptimizers	\$1,191	\$1,571	\$1,145
Average Adjusted Gross Income	\$16,578	\$3,970	\$17,457
Average Nontaxable Combat Pay	\$22,519	\$21,426	\$22,595
Use a Paid Tax Preparer	54%	57%	54%
Optimization Rate by Paid Preparer Use:			
No Paid Tax Preparer	73%	54%	74%
Average EITC Forgone Among Nonoptimizers	\$1,232	\$1,769	\$1,171
Average Adjusted Gross Income	\$16,694	\$3,699	\$17,540
Average Nontaxable Combat Pay	\$22,442	\$20,722	\$22,554
Yes Paid Tax Preparer	91%	84%	91%
Average EITC Forgone Among Nonoptimizers	\$1,083	\$1,123	\$1,077
Average Adjusted Gross Income	\$16,480	\$4,173	\$17,388
Average Nontaxable Combat Pay	\$22,584	\$21,954	\$22,631
Observations	987,454	64,399	922,968
Share	100%	6.5%	93.5%

NOTE: Monetary variables are in 2009 dollars.

Among nonoptimizers, the average amount of EITC forgone among personnel who should include their combat pay but do not is \$1,571, which is over \$400 more than the average amount forgone among personnel who should exclude their combat pay but do not. As expected, personnel who use a paid preparer are more likely to optimize their EITC. Overall, 91 percent of personnel who use a paid tax preparer optimize their EITC versus 73 percent of personnel who do not use one. Conditional on not using a paid preparer, 54 percent of personnel who should include their combat pay and 74 percent of personnel who should exclude it optimize their EITC. Among those who do not use a paid preparer, the average amount forgone among personnel whose optimal strategy is to exclude their combat pay is almost \$1,800, or nearly \$600 greater than personnel who should exclude it. Even with the use of a paid tax preparer, the optimization rate among personnel who should include nontaxable combat pay is lower than the rate among personnel who should not. In particular, 84 percent of personnel using a paid preparer who should include their combat pay optimize their EITC, which is 7 percentage points lower than the optimization rate of personnel who should exclude it.

Table 2 also demonstrates that personnel whose optimal strategy is to invoke the NCPE and who do not use a paid tax preparer are not only the least likely to optimize their EITC, but also have the lowest average AGI and nontaxable combat pay. Based on these measures, the NCPE is the least effective for personnel with the lowest income. This result could suggest that better tax guidance might need to be provided to low-income military personnel. In particular, the NCPE creates additional complexity to the EITC calculation, and this complexity is not always overcome by the use of a paid tax preparer, implying that better training of preparers on military tax issues might be necessary.

Table 3 presents the distribution of nonoptimizing personnel by the amount of EITC forgone due to not applying the NCPE optimally. These statistics are separated by whether or not the EITC is claimed. Nonoptimizing personnel could be making mistakes that prevent them from claiming the optimal EITC, or the cost of figuring out how to calculate the maximum credit could be too high relative to the amount of credit that could be claimed. The highest proportion of nonoptimizing EITC claimants (27 percent) forgoes \$250 or less, while the lowest proportion (12 percent) forgoes more than \$2,000. Among nonclaimants, the distribution of personnel by EITC forgone is heavily skewed towards the low and high end. Specifically, over 65 percent of filers and over 45 percent of nonfilers forgo \$500 or less in EITC while 21 percent of filers and 39 percent of nonfilers forgo more than \$2,000.

TABLE 3. Distribution of Nonoptimizing Personnel by Earned Income Tax Credit (EITC) Forgone, Tax Years 2005–2009

	All Nonoptimizers	EITC Claimants	Non-EITC Claimants	
			Filers	Nonfilers
EITC Forgone ≤ \$250	30%	27%	34%	22%
\$250 < EITC Forgone ≤ \$500	29%	17%	33%	25%
\$500 < EITC Forgone ≤ \$1,000	5%	21%	4%	4%
\$1,000 < EITC Forgone ≤ \$2,000	10%	23%	8%	11%
\$2,000 < EITC Forgone	26%	12%	21%	39%
Total	100%	100%	100%	100%
Observations	172,808	12,011	106,637	54,160

NOTE: Totals may not add to 100% due to rounding.

If nonoptimizing personnel located on the low end of the income distribution view the cost of determining the optimal EITC to be too high, then these personnel are not nonoptimizers in an overall sense. If personnel who could claim \$250 or less in additional EITC are reclassified as optimizers, then this would reduce the nonoptimizing population by over 50,000 and increase the overall optimization rate from 82 percent to 88 percent. However, disparities in optimization rates by optimal strategy would remain. In particular, 76 percent of personnel whose optimal strategy is to apply the NCPE optimize their EITC while 88 percent of personnel who should exclude combat pay optimize their EITC.

4.2 EITC Eligibility

Table 4 provides summary statistics by whether or not an individual would qualify for the EITC if combat pay were required to be included in EITC earned income. Taxpayers who would qualify for the EITC even if combat pay were included in EITC earned income are called Always Eligible, while taxpayers who would not qualify for the EITC if combat pay were taxable are called Newly Eligible. Always Eligible taxpayers include those with zero taxable earnings. Over 50 percent of the EITC-eligible sample, or over half a million personnel, are Newly Eligible personnel, meaning that the exclusion of combat pay increases EITC eligibility by more than twofold among the sample of personnel serving in combat zones.

The distribution of personnel by service branch is roughly the same across the two groups, while Newly Eligible personnel contain a higher proportion of reserve personnel than those who are Always Eligible. As expected, a greater proportion of Newly Eligible personnel are concentrated in the higher salaried pay grades since they can currently exclude their combat pay from EITC earned income. Sixteen percent of Always Eligible personnel are in the ME06 to ME09 pay grades, compared to 31 percent of Newly Eligible personnel. In 2009, these pay grades ranged from \$2,175.50 to \$6,863.10 per month, as reported in Table A-1. Similarly, a larger proportion of Newly Eligible personnel are officers than Always Eligible personnel (14 percent versus 1 percent). Officers could earn up to \$8,513.10 per month, or over \$100,000 in 2009, which is well above the EITC means-tested threshold.

Comparing demographic characteristics reveals that, on average, Newly Eligible personnel are older than Always Eligible personnel. Newly Eligible personnel claim an average EITC of \$1,335 while Always Eligible personnel claim an average EITC of \$2,902. Part of this difference is explained by the fact that Newly Eligible personnel are eligible for a lower maximum amount of EITC because a higher proportion file as single and a lower proportion have EITC qualifying dependents than Always Eligible personnel. In addition, a smaller percentage of Newly Eligible personnel claim the EITC (79 percent versus 90 percent), contributing to their lower optimization rate. As discussed earlier, some of these Newly Eligible personnel, specifically those who file as single, may view the cost of calculating their EITC as greater than the potential EITC they could claim.

Newly Eligible personnel have an average of \$28,658 in nontaxable combat pay, which is much larger than that of Always Eligible personnel (\$13,867). Just over half of personnel in each group use a paid preparer and about 5 percent of each group are nonfilers. By definition, Newly Eligible personnel should not apply the NCPE if they want to maximize their EITC. Less than 1 percent of these taxpayers exercise this election, while 13 percent of Always Eligible personnel use the NCPE.

TABLE 4. Summary Statistics by Always and Newly Eligible Personnel, Tax Years 2005–2009

Item	(1) Always EITC-Eligible		(2) Newly EITC-Eligible	
	Share or Mean	Standard Deviation	Share or Mean	Standard Deviation
Distribution Across Military Service Branch				
Army	63%		68%	
Air Force	13%		11%	
Marines	10%		9%	
Navy	14%		12%	
Distribution Across Military Component				
National Guard	12%		16%	
Reserves	7%		12%	
Active	81%		72%	
Distribution by Pay Grade^a				
Enlisted: Pay Grade ME01-ME03	12%		4%	
Enlisted: Pay Grade ME04	37%		20%	
Enlisted: Pay Grade ME05	34%		24%	
Enlisted: Pay Grade ME06	15%		21%	
Enlisted: Pay Grade ME07-ME09	1%		17%	
Officer	1%		14%	
Characteristics				
Age	27.97	(5.71)	32.14	(6.75)
Single	7%		26%	
Married Filing Jointly	81%		64%	
Head of Household	11%		10%	
Any EITC Qualifying Dependents	97%		60%	
Earned Income Tax Credit	\$2,902	(1,399)	\$1,335	(1,482)
Adjusted Gross Income	\$16,024	(9,094)	\$17,008	(13,180)
Nontaxable Combat Pay	\$13,867	(8,894)	\$28,658	(15,869)
Paid Preparer	56%		53%	
Claim EITC	90%		79%	
Nonfiler	6%		5%	
Use Nontaxable Combat Pay Election	13%		0%	
Optimize EITC	87%		79%	
Observations	431,590		555,777	

NOTE: Always Eligible are personnel who would be eligible for the EITC even if combat pay were required to be included in EITC earned income, while Newly Eligible are personnel who would not be eligible for the EITC if combat pay were required to be included. Monetary variables are in 2009 dollars.

^a **Monthly Basic Pay Ranges for Military Personnel by Pay Grade for 2009**

Pay Grade	Minimum (\$)	Maximum (\$)
ME01-ME03	1,294.50	1,859.70
ME04	1,827.60	2,218.50
ME05	1,993.50	2,828.40
ME06	2,175.50	3,369.90
ME07-ME09	2,515.50	6,863.10
Officer	2,655.30	8,513.10

NOTE: Source: <http://militarypay.defense.gov/Pay/basicpay.html>.

4.3 Cost Estimates

In this section, we estimate the cost of the NCPE to the Federal Government. The cost of the NCPE is part of the overall Combat Zone Tax Exclusion (CZTE) tax expenditure, which is the total amount of Federal income tax revenue forgone by making combat zone pay exempt from income taxes (Gleason and Tong, 2012). The cost of the NCPE is the sum of the difference between the EITC claimed with the inclusion of combat pay and the EITC that would have been claimed without the inclusion of combat pay among personnel who use the NCPE. Table 5 demonstrates that the tax expenditure fluctuates across years and varies between \$7.2 million in 2006 and \$15.1 million in 2009. The bulk of this expenditure comes from individuals in the ME01 through ME06 pay grades. Compared to the overall CZTE tax expenditure, the NCPE represents less than 1 percent of the total cost to the government.

TABLE 5. Cost Estimates of CZTE and NCPE Tax Expenditures, Tax Years 2005–2009

(In millions of dollars)

	2005	2006	2007	2008	2009
Total Cost of NCPE	9.3	7.2	8.0	9.6	15.1
Combat Zone Tax Exclusion	3,500	3,400	3,900	3,800	3,600

NOTE: Money amounts are in 2009 dollars.

5. Summary and Policy Implications

We have examined the relationship between nontaxable combat pay and the EITC. The evaluation of the NCPE illustrates that policies implemented through the tax code may help targeted populations, but they can also have unintended effects. While the overall optimization rate is high at 82 percent, EITC optimization rates for those who *should* use this option are lower than the rates for personnel who *should not* use this option. These differences remain when conditioning on using a paid tax preparer and become even larger when looking at the subgroup who do not use a paid tax preparer. Even though the NCPE was designed to improve outcomes for lower-income personnel, and does so most of the time, the policy makes certain personnel worse off because it adds more complexity to the EITC calculation. Since paid tax preparers also make mistakes on when to use the NCPE, better training for preparers on military tax issues might be needed. We estimate that the cost associated with the NCPE is relatively small, totaling no more than \$15.1 million, or less than 1 percent of the overall cost of the CZTE in a given year.

We also find that well over half of the EITC-eligible military population in our sample, or 555,777 out of the 987,367 personnel, is eligible for the EITC only because of the exclusion of combat pay (see Table 4). When these personnel are not stationed in a combat zone, they would generally have income levels that would exceed the threshold to qualify for the EITC, which is a tax credit targeted to lower income families. Changing the NCPE to a requirement to include combat pay in the calculation of earned income for the EITC would ensure that the lowest income military families would be able to claim the EITC and simplify the tax calculations for these families.

A more general implication of this paper is that increasing complexity in the tax code increases filing burden—leading many low-income families to rely on paid preparers in order to maximize their benefits and minimize their tax. Opportunities to simplify the tax code, particularly provisions targeting low income families should be a priority for tax policy going forward.

References

- Chetty, Raj, and Emmanuel Saez. 2013. "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." *American Economic Journal: Applied Economics*, 5(1), 1–31.
- Gleason, S., and Tong, P. 2012. "The Nontaxable Combat Pay Election and the Earned Income Tax Credit." Paper presented at the *American Economic Association Annual Conference*, (January 2012).
- Plueger, Dean. 2009. "Earned Income Tax Credit Participation Rate for Tax Year 2005," Internal Revenue Service, U.S. Department of the Treasury, *IRS Research Bulletin*, Publication 1500.
- Statistics of Income. 2011. *2009 Estimated Data Line Counts Individual Income Tax Returns*. Internal Revenue Service, U.S. Department of the Treasury, Document 12276.
- U.S. General Accounting Office. 2001. "Earned Income Tax Credit Eligibility and Participation." GAO-02-290R. U.S. GAO, Washington, D.C.
- U.S. General Accounting Office. 2004. "Military Personnel: Active Duty Compensation and Its Tax Treatment." GAO-04-721R. U.S. GAO, Washington, D.C.

Incorporating Supplemental Nutrition Assistance Program Data in Earned Income Tax Credit Administration: A Florida Case Study

*Elaine Maag, with Sara Edelstein, Devlin Hanson, Sarah Minton, Michael Pergamit,
and Caroline Ratcliffe, The Urban Institute*

Introduction

The earned income tax credit (EITC) is the most effective antipoverty program targeted to working-age households. In 2012, the credit provided \$64.1 billion to 27.8 million tax units (IRS 2012). The lion's share of EITC payments went to families with children (97 percent).¹

The IRS faces two main challenges to the efficient administration of the EITC. First, some individuals and families who are ineligible for the EITC claim the credit erroneously. Second and conversely, some individuals and families who are eligible for the EITC fail to claim the credit. The IRS works to correct both of these errors.

Using data from the IRS's National Research Program (NRP) from Tax Year 2009, the IRS estimates that for Fiscal Year 2013, between 22.1 percent and 25.9 percent of total EITC program payments were overclaims (U.S. Department of the Treasury (2013)). On tax returns claiming an EITC between 2006 and 2008, determining whether a child met the complex "qualifying child" rules was the largest source of error (IRS (2014)). Determining whether a child passed the residency test was a source of error on 75 percent of returns with qualifying child errors. The Office of Management and Budget has identified the EITC as having the highest improper payment rate and second-highest improper payment amount among 13 "high-error" programs.²

Unlike some information necessary for calculating taxes (e.g., wages, interest payments, and mortgage interest), the IRS does not receive information from a third party verifying where and with whom a child resides; yet residency is an important element in determining whether a child meets the EITC test of being a qualifying child. To reduce errors associated with the qualifying child test, Congress could simplify the EITC eligibility criteria to remove the residency test or the IRS could develop third-party data that could verify EITC claims. We explore the latter option by using Florida SNAP data matched to IRS data to determine whether such data could be used to verify whether a child meets the residency test.³

Our analysis finds that half of all tax units in Florida that claimed the EITC have at least one member who appears in the SNAP data. Although SNAP data contain some information on household structure, the data are only suggestive about EITC eligibility. They are not definitive enough to warrant delaying a refund. However, these data might help identify returns for audit because of evidence that a child might not meet the residency test. The data may also help determine whether someone claiming the childless EITC has a qualifying child and is thus ineligible to claim the much smaller "childless" EITC, even if he or she does not claim the EITC for workers with children. Finally, SNAP data may help the IRS target outreach efforts to potentially eligible nonclaimants who failed to file a tax return.

¹ Tax Policy Center. 2013. "Table T-13-0221: Tax Benefit of the Earned Income Tax Credit." Washington, DC: Urban Institute. <http://www.taxpolicycenter.org/numbers/displayatab.cfm?DocID=3980&topic2ID=60&topic3ID=65&DocTypeID=>

² See "High Error Programs," Payment Accuracy, accessed July 1, 2015, <http://www.paymentaccuracy.gov/high-priority-programs>. The Office of Management and Budget defines high-error programs as "those programs that reported roughly \$750 million or more in improper payments in a given year, did not report an error amount in the current reporting year but previously reported an error amount over the threshold, or have not yet established a program error rate and have measured components that were above the threshold."

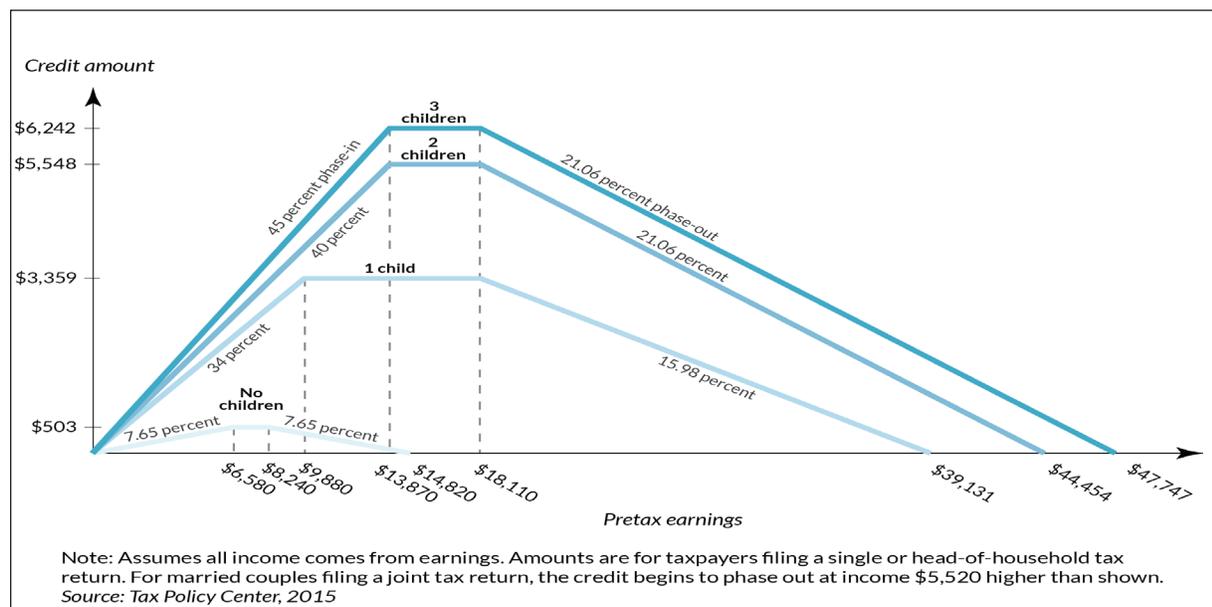
³ This report is part of a larger project documented in Pergamit, *et al.* (2014); a similar shortened version of this study was published as Maag, *et al.* (2015).

Understanding the EITC and SNAP

The Earned Income Tax Credit

The EITC subsidizes low-income working families. The credit equals a fixed percentage of earnings from the first dollar of earnings until the credit reaches its maximum; both the percentage and the maximum credit depend on the number of children in the family. The credit then stays flat at the maximum as earnings continue to rise, but earnings eventually reach a phase-out range. From that point, the credit falls with each additional dollar of income until it disappears entirely (Figure 1). The phaseout begins at a higher income for married couples than for individuals. The credit is fully refundable: any excess beyond a family's income tax liability is paid as a tax refund. In Tax Year 2015, the maximum credit ranges from \$503 for tax filers with no children to \$6,242 for families with three or more children. All credit thresholds are adjusted annually for inflation.

FIGURE 1. Earned Income Tax Credit, 2015



SOURCE: Tax Policy Center (2015).

NOTE: Assumes all income comes from earnings. Amounts are for taxpayers filing a single or head-of-household tax return. For married couples filing a joint tax return, the credit begins to phase out at income \$5,520 higher than shown.

Complex characteristics determine eligibility for and size of the EITC, including income (both earned and unearned), demographic characteristics (citizenship status, marital status, age, relationship), and residency. Some ineligible EITC claimants receive the tax credit (U.S. Department of the Treasury (2011)) and some eligible people fail to claim it. EITC-eligible individuals with incomes below the tax filing threshold are not required to file tax returns for income tax purposes, but by not filing they miss the opportunity to receive the EITC. Consider a married couple with two children. If all of their gross income came from earnings, they could earn of up to \$19,000 before being required to file a tax return. However, failing to do so would mean they would forgo an EITC of over \$5,000. Analysts estimate that between 70 and 89 percent of people eligible for the EITC actually receive it (Blumenthal, Erard, and Ho (2005), Olson (2014), Plueger (2009), Scholz (1994), White (2002)).

Most information required for determining EITC eligibility and the credit amount applies to all people claiming the EITC. This includes having a Social Security number valid for work, filing status, and income and earnings. Families with children must also establish whether each child is a qualifying child by passing four tests:

- (1) **Relationship test:** A qualifying child must have a specific relationship to the tax filer (child or step-child—whether by blood or adoption, foster child, sibling, half-sibling, stepsibling, or descendant of any of these).
- (2) **Residency test:** A qualifying child must live with the taxpayer in the United States for more than half of the year.
- (3) **Age test:** The child must be younger than the taxpayer (or spouse) and be: (a) under age 19; (b) between the ages of 19 and 23 and attending school full time during at least five months of the year; or (c) totally disabled.
- (4) **Joint return test:** The child cannot file a joint return unless it was filed only to claim a refund of withheld taxes.

Children sometimes meet the requirements of being a qualifying child for more than one person, and sorting out who ought to claim the child can be confusing. A child may serve as a qualifying child for only one tax filer, even if he or she meets the requirements for multiple tax filers.⁴

Tax filers claiming the childless EITC must also meet an age requirement and certain other restrictions, including not having a qualifying child (even if they do not claim that child for tax purposes).

Currently, the IRS does not receive third-party information that allows verification that a child and adult live together in the U.S. for more than half the year, whether a 19- to 23-year-old is in school full time during at least five months of the year, a child's relationship to an adult who is not the child's parent, or a person's disability status. In addition, only incomplete information is available to verify self-employment income. These are all important factors in calculating a person's EITC.

The Supplemental Nutrition Assistance Program

SNAP provides low-income families with funds for food purchases via electronic cards. Benefits are based primarily on household makeup (people who purchase and prepare meals together) and income. It is likely that many people eligible for SNAP benefits are also eligible for the EITC.

The Food and Nutrition Service in the U.S. Department of Agriculture runs the SNAP program. The Federal Government pays for all benefits and splits the cost of program administration with the States. To receive SNAP benefits, a household must have gross income (income before deductions, which vary by State) less than or equal to 130 percent of the Federal poverty level (FPL), and net income (gross income minus deductions) no more than 100 percent of FPL (table 1). Households with an elderly or disabled person need to meet only the net income test, though these households are less likely to be part of the EITC population because their likelihood of employment is lower and they are less likely to have qualifying children.

TABLE 1. SNAP Income Test, October 1, 2014–September 30, 2015

Household size	Gross monthly income (130 percent of Federal poverty level) (\$)	Net monthly income (100 percent of Federal poverty level) (\$)
1	1,265	973
2	1,705	1,311
3	2,144	1,650
Each additional member	440	339

SOURCE: "Supplemental Nutrition Assistance Program (SNAP)," USDA Food and Nutrition Service, accessed July 16, 2015, <http://www.fns.usda.gov/snap/eligibility#Income>.

NOTE: SNAP = Supplemental Nutrition Assistance Program.

⁴ If the child meets the requirements to serve as a qualifying child for more than one tax filer, the IRS invokes tiebreaker rules. The tiebreaker rules are complicated and have changed frequently. In general, a parent's claim takes precedence. If a child is the qualifying child of two unmarried parents, the parents may choose who may claim the child; however, if both parents want to claim the child, then the parent with whom the child lived the longest, or if equal, the parent with the higher adjusted gross income (AGI), may claim the child. If no parent claims the child, the person with the highest AGI among those for whom the child is a qualifying child may claim the child instead of the parents. In cases where a parent could claim the child but does not do so, the person claiming the child must have AGI higher than the parent's AGI.

Most households must have counted assets of less than \$2,000 to qualify for SNAP benefits. Cash, bank accounts, and stocks or bonds are examples of counted assets; excluded are family homes, business properties, retirement accounts, education accounts, and in all but three States, the value of one or all vehicles. (EITC refunds can be saved for 12 months after receipt before counted toward the SNAP asset limit.) Relative to the EITC, SNAP asset limits are more restrictive because the EITC rules limit only the level of income received from investments, not the total value of assets. Households in which all members receive Supplemental Security Income, Temporary Assistance to Needy Families, or General Assistance—programs that have far more restrictive eligibility rules—do not need to pass the income and asset tests because they are automatically eligible for SNAP. About one-quarter of SNAP caseloads in 2010 received these types of benefits (USDA (2011)).

Broad-based categorical eligibility (BBCE), a policy that 42 States and the District of Columbia have implemented as of 2013, further expands eligibility for SNAP, though usually not beyond the scope of the EITC. Under BBCE policies, households that are authorized to receive a Temporary Assistance for Needy Families Maintenance-of-Effort-funded noncash benefit (which usually takes the form of a brochure targeted to virtually all SNAP applicants) are automatically eligible for SNAP. Gross income limits under BBCE range from 130 percent to 200 percent of FPL; 14 States use the latter limit. Some States retain the Federal net income limit of 100 percent of FPL, but most do not. However, because benefit calculations are still based on net income, some households that have broad-based categorical eligibility may not actually be eligible to receive a SNAP payment.

Determining exactly how many families are eligible due solely to BBCE is not possible, because the asset information on families in States with these policies often is not collected. In 2008, among the eight States that had broad-based programs with gross income limits over 130 percent of FPL, an estimated 6 percent of households eligible through BBCE policies would have been ineligible without them (Trippe and Gillooly (2010)). Depending on the generosity of expanded SNAP gross income limits established through more newly implemented BBCE policies, this percentage may have grown.

Federal guidelines require a SNAP recertification period of no more than 12 months, or 24 months if all household members are elderly or disabled. In 2010, 12 months was the average period for all households across States (Eslami, Filion, and Strayer (2011)). This represents a trend away from much shorter recertification periods. In 2000, some 19 percent of households recertified every 3 months or less; in 2009 only 1 percent of households did so (Klerman and Danielson (2011)). During the recertification interviews, SNAP beneficiaries must update information used to determine eligibility, including household composition.

In 2010, SNAP served 18.6 million households, and 49 percent of those households contained children. Of SNAP households with children, 48 percent have earned income (Eslami, Filion, and Strayer (2011)), a necessary element to being eligible for the EITC.

Intersection of SNAP and EITC in Florida Administrative Data

Not all EITC recipients will be eligible for or receive SNAP, and not all SNAP recipients will be eligible for the EITC. Importantly, EITC eligibility is based on an annual measure of earnings, while SNAP benefits are based on monthly earnings. Assuming annual earnings are spread out evenly across all months of the year, the EITC is available to individuals at higher income levels than SNAP benefits. It is possible to have low earnings in some months and qualify for SNAP but move to a higher-paying job later on and lose eligibility for both SNAP in those months and the EITC for the entire year. Although an individual (or spouse, if married) must work and have a valid Social Security number to be eligible for the EITC,⁵ SNAP benefits can be available to people without earnings (particularly people with disabilities or others exempt from work requirements). Also, some States allow individuals without valid Social Security numbers to receive SNAP benefits.

⁵ In the case of a married couple, both spouses must have a valid Social Security number for their work to be eligible for the EITC.

In 2010, almost 4.9 million people in Florida were in a tax unit claiming the EITC (Table 2). About half (2.4 million) of the people in a tax unit claiming the EITC also received SNAP.⁶ These people were in 2.2 million tax units. Half of those tax units (1.1 million) contained at least one person who also received SNAP. Most tax units claiming the EITC found in SNAP units reported at least one qualifying child on their tax return. About 182,000 tax units with at least one person in the Florida SNAP data had no qualifying child.⁷

This study relies on SNAP benefit data from Florida in 2010, a period when the overlap between the two programs was likely unusually high. As the economy recovered, many SNAP beneficiaries probably continued to be eligible for the EITC, but the total number of SNAP beneficiaries declined.

TABLE 2. EITC Claimants Found in Florida SNAP Data, 2010

	Number	Percentage
Individuals in Tax Units Claiming EITC	4,890,219	100.0
• Those in State SNAP data	2,444,276	50.0
• Not in State SNAP data	2,445,943	50.0
Number of Tax Units Claiming EITC	2,168,369	100.0
• Matched tax units (tax units with at least one person receiving SNAP benefits)	1,087,308	50.1
<i>Claiming children for EITC purposes</i>	905,009	41.7
All members found in State data	528,257	24.4
All tax filer(s) and at least one child found in State data, but not all children	42,499	2.0
All tax filer(s), but none of the children found in State data	100,527	4.6
No tax filer(s), but all of the children found in the State data	208,449	9.6
At least one tax filer but not all tax filer(s) found in State data	25,277	1.2
<i>Claiming childless EITC</i>	182,299	8.4
At least one tax filer found in State data	182,299	8.4
• Unmatched tax units (tax units with no people identified in SNAP data)	1,081,061	49.9

SOURCE: IRS data matched to Florida SNAP benefit receipt data

NOTE: EITC = earned income tax credit; SNAP = Supplemental Nutrition Assistance Program

We test various aspects of EITC eligibility based on which people in the tax unit appear in the Florida SNAP data. For some tax units, we find all members of the tax unit in the State data (24.4 percent). For tax units with all members found in the SNAP data, we analyze both the EITC residency and relationship tests. We can test whether all children in the tax unit appear to be related by one of the qualifying relationships to the adults in the tax unit and observe how many months the children appear to live with the adults in the tax unit. For the 2 percent of tax units in which we observe only a portion of the children, we can test these same things, but only for a portion of the tax unit. For the 4.6 percent of tax units in which we observe adults but none of the children, we have no SNAP data on the relationship of the children being claimed in the tax unit, but we can infer that children claimed by the tax unit are unlikely to have lived in the tax unit long enough if we observe the adults for more than half the year. Finally, in the 9.6 percent of tax units in which we observe children in the SNAP data but not the person claiming them for tax purposes, we can estimate how long these children appear to have lived with someone other than the tax unit claiming them during the year. In cases where they lived apart from the person who claimed them as a qualifying child for more than half the year, they are unlikely to have passed the residency test.

⁶ A very small percentage (less than 1 percent) of people observed in the Florida SNAP data did not actually receive SNAP benefits. These individuals lived with the SNAP beneficiary but were not part of the assistance unit. For purposes of this report, we include in our analysis everyone for whom we have information, even if they did not actually receive benefits.

⁷ In a small number of cases, Social Security numbers recorded in SNAP data were invalid and others were not recorded. We exclude these cases from this analysis.

Our final analysis group comprises individuals who appear in the State data but not the IRS data. We analyze these cases to see if outreach efforts aimed at this population would be well targeted.

Using SNAP Benefit Data To Verify Qualifying Children in EITC Claims: Relationship and Residency

Using Florida SNAP data matched to IRS data, we analyzed the relationship and residency of qualifying children. In total, over a quarter million Florida tax returns appear to have a child fail the residency test. This represents 11.7 percent of all Florida tax returns with an EITC claim and 15.7 percent of all Florida tax returns claiming the EITC with children. We identify failures as children whom we observe with a household other than the one claiming the child for EITC purposes for at least 6 months, or as a taxpayer who claims a child and the child does not appear to be in the SNAP unit containing the taxpayer for at least 6 months of the year. However, the SNAP data do not record whether a child is present in every month, so we assume that if a child is observed in two periods with an adult, they remain with the adult during the intervening months.

If an adult has a qualifying child for the EITC, regardless of whether they claim that qualifying child, they are ineligible to receive the childless EITC. Our analyses of Florida SNAP data indicate approximately one-eighth (12 percent) of all tax units in Florida that claimed the childless EITC in 2010 may have had an improper claim, based on having a qualifying child.

Finally, our analysis indicates there may be some people who receive SNAP benefits who appear eligible for the EITC but do not file a tax return.

Tax Returns for Families Claiming the EITC for Workers with Children

Tax Units with All Adults and Children Found in the State Data

There are 528,257 fully matched tax units in which all adults claiming an EITC and all children claimed for the EITC are found in the State data. All children in almost all fully matched tax units claiming the EITC for families with children have relationships reported that are consistent with EITC qualifying relationships (99.1 percent; Table 3).

Determining whether children in these fully matched units meet the residency test is less straightforward than for the relationship test, largely because we must make assumptions about where a child lives during months that are not directly reported to the SNAP office.

To pass the EITC residency test, a claimant and his or her qualifying child must live together for more than half of the year. Ideally, we would verify residency by observing an EITC claimant and his or her qualifying child together in 6 months of SNAP data. Conversely, if we observed an EITC claimant and the child he or she claims in different households for at least 6 months, we could consider the pair to have failed the residency test.

Florida SNAP benefit data do not provide enough information to determine whether a claimant passes or fails the residency test. Less than 1 percent of individuals appear in a recertification in 6 different months and more than half appear in a recertification only once in 2010 (Pergamit, *et al.* (2014)). We do not know the frequency of reporting residency to the State. Although many families received benefits for the full 12 months of 2010, we analyzed individual case records in Florida and learned that most SNAP recipients appeared in a recertification once or twice in 2010, which would be the only direct reports of household composition available in the Florida SNAP data. Our analysis assumes that a child who appears in a SNAP unit remains in that unit until a change is reported to the SNAP administrators, consistent with how benefits are calculated.

We assess length of residency by analyzing how many months a child is considered to be in a SNAP unit during the year. Each case presumably recertifies one or two times a year based on Florida's 6-month recertification period; most families will not recertify more than twice a year. Once an application has been submitted, the household structure is assumed by Florida program administrators to remain unchanged until the next recertification, when it is updated based on new information provided by the recipient. SNAP rules in Florida do not require beneficiaries to report household changes except at the time of recertification. If an adult and

child appear together in the SNAP data at least twice during the year (which would mean they received benefits in at least 2 months), with at least 4 months between observations (so covering a 6-month period) and up to 10 months between those two appearances, we consider them to have passed the residency test. In cases in which an adult and child appear together in the SNAP data for fewer than 6 months, we consider there to be insufficient information for determining whether the residency test was met.

Applying these criteria, we find that 76.9 percent of fully matched tax units pass the residency test for all children being claimed for the EITC (which represents 19 percent of all tax units claiming the EITC). In 2.1 percent of these tax units, some children pass the residency test (thus, at the very least the EITC claimants would still be eligible for an EITC for families with one child) and in the remaining 20.5 percent of these tax units, no children claimed for the EITC pass the residency test. In most cases where not all children pass the residency test, the reason is because there is insufficient evidence to evaluate the test for at least some children. In only 0.5 percent of cases is there sufficient evidence to conclude that all children fail the residency test. In another 0.3 percent, at least one child fails, but some either pass or are inconclusive.

TABLE 3. Analysis of the Qualifying Child Tests for Tax Units with All Members in the Florida SNAP Data, 2010

Tax units	528,257
Relationship test ^a	
<i>Full tax unit passes</i>	99.1%
<i>At least one child passes</i>	0.6%
At least one child fails	0.4%
At least one child has insufficient information	0.2%
<i>No children pass</i>	0.3%
At least one child fails	0.0%
All children have insufficient information	0.1%
All children fail	0.2%
Residency test ^b	
<i>Full tax unit passes</i>	76.9%
<i>At least one child passes</i>	2.1%
At least one child fails	0.2%
At least one child has insufficient information	1.9%
<i>No children pass</i>	20.5%
At least one child fails	0.1%
All children have insufficient information	20.5%
<i>All children fail</i>	0.5%

SOURCE: IRS data matched to Florida SNAP benefit receipt data.

NOTE: SNAP = Supplemental Nutrition Assistance Program. Sample composed of tax units claiming the EITC in the IRS data in which the primary tax filer, secondary tax filer (if applicable), and children claimed for the EITC are matched to individuals in the Florida SNAP data.

^a Relationship test: "Fail" if the tax filer(s) and the child are not qualifying relatives; "Have insufficient information" if the tax filer(s) and the child have an ambiguous or missing relationship; "Pass" if the tax filer(s) and the child are qualifying relatives.

^b Residency test: "Fail" if either the tax filer(s) or the child receives benefits in a case without the other covering a 6-month period with fewer than 10 months between benefit receipts; "Have insufficient information" if the tax filer(s) and the child receive benefits in the same case but the benefit receipt does not cover a 6-month period or there are more than 10 months between benefit receipts; "Pass" if the tax filer(s) and the child receive benefits in the same case over a 6-month period with no more than 10 months between benefit receipts.

Tax Units with Adults Found in the State Data, But Not All Children Found in the State Data

There are 143,026 tax returns in which not all children claimed for the EITC are found in the State data. In 42,499 we observe at least one child claimed on the return in the State data (but not in the same household as claimed for EITC purposes), and in the remaining 100,527 we observe no children that were claimed for the EITC in the State data (Table 4). We apply the same residency tests to these returns as was applied to fully matched tax units to see if State data can verify either the existence of at least one child for the requisite period

or supply evidence that it is unlikely the children claimed on the tax return but not found in State data are actually in the tax unit. We find some evidence that in 54.8 percent of the 42,499 tax units in which we observe an adult that claimed a child for the EITC and at least one child who was claimed is observed in the State data, the child claimed for the EITC appeared to live for at least 6 months in a different SNAP unit than the one the adult claiming them for the EITC lived in. Up to 16 percent more tax units in this group claimed a child for the EITC who did not live with them for at least some months, according to their SNAP reports. In cases where no children were found in the State data, we find evidence of adults who claimed a child for the EITC not reporting to SNAP offices at least one child in 1.9 percent of tax returns and reporting to SNAP offices no children eligible for the EITC in 57.9 percent of cases.

TABLE 4. Analysis of the Relationship and Residency Tests for Tax Units with the Tax Filer(s) in the Florida SNAP Data, but at Least One Child Not in the Florida SNAP Data, 2010

	At least one child found in the State data, but not all	No children found in the State data
Tax units	42,499	100,527
Relationship test ^a		
<i>At least one child passes</i>	99.1%	
At least one child fails	0.2%	
At least one child has insufficient information	98.9%	
<i>No children pass</i>	0.8%	
At least one child fails	0.5%	
All children have insufficient information	0.3%	
<i>All children fail</i>	0.0%	
Residency test ^b		
<i>At least one child passes</i>	70.8%	
At least one child fails	54.8%	
At least one child has insufficient information	16.0%	
<i>No children pass</i>	29.0%	42.1%
At least one child fails	2.4%	1.9%
All children have insufficient information	26.6%	40.2%
<i>All children fail</i>	0.2%	57.9%

SOURCE: IRS data matched to Florida SNAP benefit receipt data.

NOTE: SNAP = Supplemental Nutrition Assistance Program. Sample composed of tax units claiming the EITC in the IRS data in which the primary and secondary (if applicable) tax filers are matched to individuals in the Florida SNAP data, but at least one of the qualifying children is *not* matched to an individual in the Florida SNAP data.

^aRelationship test: "Fail" if the tax filer(s) and child are not qualifying relatives; "Have insufficient information" if the tax filer(s) and child have an ambiguous or missing relationship; "Pass" if the tax filer(s) and child are qualifying relatives.

^bResidency test: "Fail" if either the tax filer(s) or the child receives benefits in a case without the other covering a 6-month period with less than 10 months between benefit receipts; "Have insufficient information" if the tax filer(s) and child receive benefits in the same case but the benefit receipt does not cover a 6-month period or there are more than 10 months between benefit receipts. For a child not in the State data, the child is considered to have insufficient information if the child has a missing Social Security number or was born in 2010, regardless of the length of benefit receipt. "Pass" if the tax filer(s) and child receive benefits in the same case over a 6-month period with no more than 10 months between benefit receipts.

Tax Units in Which Children Are Found in the State Data, but Adults Are Not Found in the State Data

The final group of tax units on which we perform the residency test are the 208,499 units that claim the EITC in which none of the adults claiming the children are found in the Florida SNAP data but some or all of the children on these returns are found in the data (representing 9.6 percent of all tax returns claiming the EITC; Table 5). In 60.8 percent of these tax units, we observe at least one of the children claimed as a qualifying child in the Florida SNAP data who, for at least 6 months, lives with an adult other than the one (or two, if married)

claiming them for EITC purposes. In another 21.1 percent of these cases, we observe the children living with an adult other than the one claiming them for fewer than 6 months (insufficient evidence).

TABLE 5. Analysis of the Residency Tests for Tax Units with Primary Tax Filers Not in the Florida SNAP Data Claiming Children in the Florida SNAP Data, 2010

Tax units	208,449
Residency test ^a	
<i>No children pass</i>	39.3%
At least one child fails	18.2%
All children have insufficient information	21.1%
<i>All children fail</i>	60.8%

SOURCE: IRS data matched to Florida SNAP benefit receipt data.

NOTE: SNAP = Supplemental Nutrition Assistance Program. Sample composed of tax units claiming the EITC in the IRS data in which the primary (and secondary if applicable) are not matched to any individual in the Florida SNAP data and at least one of the children they claimed for the EITC is matched to an individual in the SNAP data.

^aResidency test: "Fail" if the child receives benefits in a case without the tax filer(s) covering a 6-month period with fewer than 10 months between benefit receipts; "Have insufficient information" if the child receives benefits in a case without the tax filer(s) but the benefit receipt does not cover a 6-month period or there are more than 10 months between benefit receipts.

Tax Returns for Families Claiming the EITC for Workers Without Children

A tax filer with a qualifying child is not permitted to file a childless EITC claim. This can happen in multi-generational households in which a grandmother claims a grandchild for the EITC and a parent attempts to claim the childless EITC, when both are eligible to claim the child. It can also happen if unmarried parents live together with their child. In either case, only one adult may claim the EITC for workers with children and neither may claim the childless EITC.

Of the 2.2 million tax returns claiming the EITC in Florida in 2010, we observe 182,299 claiming the childless EITC, representing 8.4 percent of all tax units in Florida claiming the EITC (Table 6). We test these tax units to see whether the claimant appears to have a qualifying child. On 77 percent of these returns, we find no evidence that the EITC claimant has a qualifying child. In 12 percent of the returns in this group that we were able to match, however, we find evidence that the childless EITC claimant has a qualifying child, which would make them ineligible to receive the childless EITC. This represents approximately 1 percent of all tax units that claimed the EITC in 2010. Note that in 74 percent of the cases wherein we believe the childless EITC claimant has a qualifying child, that qualifying child is claimed on another return (not shown).

TABLE 6. Analysis of Childless EITC Claims in Which the Primary Tax Filer Is Matched in the Florida SNAP Data, 2010

Tax units	182,299
Has qualifying child ^a	
Some evidence	12.4%
Insufficient information	10.3%
No evidence	77.3%

SOURCE: IRS data matched to Florida SNAP benefit receipt data.

NOTE: EITC = earned income tax credit; SNAP = Supplemental Nutrition Assistance Program. Sample composed of tax units claiming the EITC in the IRS data in which the primary tax filer (and secondary if applicable) are matched to individuals in the Florida SNAP data and claim EITC without any children.

^aQualifying child: "Some evidence" if the primary tax filer is in a case with at least one child who was under 23 or disabled, younger than the primary tax filer, was a qualifying relative of the primary tax filer and was with the primary tax filer over at least a 6-month period with fewer than 10 months between sightings in the State data; "Insufficient information" if the primary tax filer is in a case with at least one child but the child does not meet at least one of the above conditions listed under "some evidence"; "No evidence" if the primary tax filer was not in a case with any children.

Nonclaimants and Nonfilers: Outreach

In some cases, we observe individuals in the Florida SNAP data who do not appear in the IRS data. To the extent that these individuals are eligible for the EITC and do not receive it, they may be candidates for IRS outreach efforts. Evidence of this in the Florida SNAP data is rare.

We observe 731,426 individuals in the SNAP data that have at least one child in the SNAP unit but do not claim the EITC. Of these, 156,518 do not file a tax return (Table 7). We do not attempt to find SNAP claimants without children who might be eligible for the EITC, given the very small income range over which someone can be eligible for the childless EITC. The income information in the SNAP data we had was not generally of high enough quality to determine whether a childless individual would qualify for an EITC.

Among the 574,908 individuals with children who file a tax return but do not claim the EITC, 121,976 pass the household level tests to be eligible for the EITC.⁸ Only a small fraction of these, 14,326 returns, have children in the SNAP unit that appear to meet the rules to be an EITC qualifying child. However, almost half of these filed returns that appear to have an eligible child (6,158), have either a recertification indicator in the IRS data, meaning they had a previous EITC claim disallowed and must recertify eligibility to receive the EITC, or IRS data indicate that the tax unit has a special processing code indicating the unit is either ineligible or does not want the EITC.

For the 156,518 people who do not appear in the tax data, 3,397 meet the household level tests as well as the qualifying child tests. In total, that leaves 11,565 people who may benefit from IRS outreach efforts aimed at notifying people of potential eligibility for the EITC.

TABLE 7. Analysis of the Potential EITC Claimants in the Florida SNAP Data, 2010

	Filed a tax return	Did not file a tax return
Found in SNAP data with child	574,908	156,518
Fails ^a	97.5%	97.8%
Passes all household and individual tests ^b	2.5%	2.2%
No recertification indicator ^c	1.4%	2.2%

SOURCE: IRS data matched to Florida SNAP benefit receipt data.

NOTE: EITC = earned income tax credit; SNAP = Supplemental Nutrition Assistance Program. Sample composed of either primary tax filers matched in the SNAP data that do not claim the EITC and have earnings and are citizens, or adults in the SNAP data that do not file a tax return.

^aIf the potential claimant is not a citizen, has no earnings, or is the qualifying child of another person according to the State data. The potential claimant must not be alone in a case or in a case with a child who does not meet all of the following criteria: is younger than the potential claimant, received benefits in the same case for a period covering at least 6 months with less than 6 months between benefit receipt and the child is not already claimed for the EITC.

^bIf the potential claimant meets all of the following criteria: is a citizen, has earnings, is not the qualifying child of another person and is in a case with at least one child who is younger than the potential claimant, is younger than 18, is a qualifying relative, is not already claimed for the EITC and received benefits in the same case for a period covering at least 6 months with less than 6 months between benefit receipts.

^cA recertification indicator is a note on the IRS data file that indicates a person was denied eligibility in a previous year and must recertify eligibility to receive the EITC.

Discussion

The EITC provides substantial assistance to low-income workers, primarily those with children. To serve as a qualifying child for the EITC, a child must live with the tax filer for more than 6 months of the year. IRS compliance studies indicate that this residency test is a large source of errors found on returns claiming the EITC. To date, the IRS has no reliable source of third-party information on where a child lives, though transfer

⁸ All adults in the tax unit are citizens with valid Social Security numbers with earnings that appear to be beneath the EITC qualifying threshold. No adult in the unit has been claimed as a qualifying child on another return, the return that has been filed does not use the status "married filing separately," and they do not file a tax form indicating they have investment income in excess of the allowable limit.

programs collect similar data. In addition, some people eligible for the EITC each year fail to get the credit, and information collected by transfer programs could help identify at least some of these nonclaimants. Transfer program information may also help determine whether someone erroneously claims the childless EITC while having a qualifying child; having such a child makes the person ineligible for the childless EITC, regardless of whether he or she claims that child.

Data from SNAP, which covers a relatively large swath of the EITC eligible population, may help the IRS. The broad definition of a SNAP assistance unit covers most household members, unlike other transfer programs, making it potentially helpful in informing EITC relationship and residency tests. SNAP income limits in many States are substantially lower than the EITC income and earnings limits, so certainly not all people claiming the EITC would be eligible for SNAP. Conversely, all EITC recipients must have a worker in the tax unit, which is not true of all SNAP recipients. The overlap between the EITC population and SNAP population should be better in States that use broad-based categorical eligibility for SNAP, but even then SNAP still fails to include all potential EITC recipients.

State benefit agencies collect eligibility information primarily at the initial program application and at benefit recertification interviews. Over the past decade, State benefit programs have begun allowing longer periods between recertifications, especially for certain subpopulations (such as disabled individuals). Generally, recertifications in SNAP occur every 12 months. The changes States require households to report before recertification vary, and the trend has moved toward reduced reporting requirements. In Florida, for example, SNAP participants are not required to report household composition changes before recertification. Consequently, household composition is known with most confidence at recertification points. This analysis assumed that if someone in the SNAP data reported living with the same people at two different times, then they lived with those same people in all intervening time periods. This assumption is not reliable enough to be used for processing under the IRS's math error (initial return checking) authority, but could possibly be useful as a filter for selecting cases to audit.

We find some evidence that tax filers may claim children who have not lived with them for more than half of the tax year (which we count as 6 full months), assuming our assumption about household stability is correct. We observe cases in which: (1) an adult is seen in the benefit data for at least 6 months without a child they claimed as a qualifying child; and (2) a child is seen in the benefit data for at least 6 months with an adult other than the one who claimed him or her as a qualifying child. In cases in which we observe the tax filer in the State data without at least one of the children he or she claimed for the EITC, the taxpayer receives SNAP benefits without that child for at least 6 months of the year in 60 percent of the tax units, which we infer to mean the child fails the residency test. This represents 4 percent of all EITC claims in Florida. In cases where we observe a child in the SNAP data without the adult who claimed them for the EITC, in 79 percent of tax units, at least one child being claimed for the EITC appeared to live with an adult other than the one claiming the child for at least 6 months of the year. Those units represent 7.6 percent of all tax units in Florida claiming the EITC.

The IRS collects information on each tax unit's home address but does not collect information on relationships of all household members at the same address. Within a tax unit, adults report whether children are dependents and qualifying children for the EITC, but the IRS knows nothing about how children living in a household relate to other household members in different tax units. This makes it impossible for the IRS to effectively administer one of the tests of eligibility for the childless EITC. That is, individuals are ineligible for the childless EITC if they have a qualifying child (IRS 2013). The rule is used to prevent a situation where one person eligible to claim a child for the EITC does so and another person eligible to claim that same child (but filing a separate tax return) claims the childless EITC.

This rule can be difficult to interpret. In the instructions for claiming the EITC, individuals follow the eligibility path for people with and without EITC-qualifying children. A potential error may occur if people assume that if they are not claiming a qualifying child, they should follow the eligibility guidelines for families without qualifying children.

We find about 182,000 cases in the Florida SNAP data in which a tax unit claims the childless EITC. Based on our assessment of the residency and relationship tests, 12 percent of these cases appear to have a qualifying child and thus are not eligible to claim a childless EITC. In 75 percent of these ineligible cases, the child we

identify as being a qualifying child of the childless EITC claimant is actually claimed as a qualifying child by another person in the IRS data.

It is possible that we identified someone as a qualifying child of a childless EITC claimant who does not actually meet the tests of eligibility. This could happen if periods of residency we assume to be stable are actually periods during which the person claiming the childless EITC is moving in and out of the household; those changes are not reflected in the SNAP data since household composition is carried forward from past months rather than updated to reflect any changes that might have occurred.

We do not know to what extent these potentially erroneous childless EITC claims in the SNAP data would be reflected in the larger group of people claiming the childless EITC. If SNAP households tend to be more complex than other households containing people claiming the childless EITC, then the rate of error may be higher in SNAP households than the more broadly eligible population.

Finally, although we find many childless EITC claimants in the SNAP data that may have qualifying children, the dollar amount associated with a childless EITC claim is small relative to that available to families with children. The benefits from pursuing errant childless EITC claims would therefore be small.

In addition to identifying possible improper EITC claims, we have examined whether a match of tax return data might identify individuals or married couples who are eligible for the EITC but either did not claim the EITC or did not file a tax return. Our analysis indicates that the State data do not provide additional information that would help the IRS conduct outreach to eligible nonclaimants who file a tax return. For the most part, tax filers who do not claim the EITC but appear in the State data do not appear to be eligible; for those who appear eligible based on State information, IRS records either have additional information indicating ineligibility or the IRS contacted these tax filers and informed them about potential EITC eligibility through postfiling notices.⁹

Eligible nonclaimants may exist among individuals who did not file a tax return. A large percentage of those who did not file a return do not appear to have wages according to the State data, leaving only 2 percent of nonfilers appearing to be eligible. However, we have noted our concerns about the quality of earnings in the State data sets. If we apply only the qualifying child tests, half of all the non-tax filers appear to have a qualifying child. Among that group, some may not be eligible for the EITC for other reasons. Further analysis could narrow the range of possible eligible nonclaimants.

Congress could improve the EITC by simplifying the program rules or by creating a “worker credit” that is available to all low-income workers, regardless of whether they have children. These changes could preserve the EITC’s work-incentive and anti-poverty benefits, while reducing errors and making the tax credit easier for workers to claim. Examples of this type of reform have been developed in Maag (2015) and proposed by Bipartisan Policy Center (2010). However, such a proposal would represent a major change that would either substantially increase the cost of the program or require cuts in benefits for many current beneficiaries.

References

- Bipartisan Policy Center. 2010. “Restoring America’s Future,” Washington, DC: Bipartisan Policy Center.
- Blumenthal, Marsha; Brian Erard; and Chih-Chin Ho. 2005. “Participation and Compliance with the Earned Income Tax Credit.” *National Tax Journal* 58 (2): 189–213.
- Eslami, Esa; Kai Filion; and Mark Strayer. 2011. “Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2010.” SNAP Report No. SNAP-11-CHAR. Washington, DC: U.S. Department of Agriculture, Food and Nutrition Service, Office of Research and Analysis.
- Internal Revenue Service. 2012. *Statistics of Income—2012 Individual Income Tax Returns*. Publication 1304. Washington, DC: Internal Revenue Service. <http://www.irs.gov/pub/irs-soi/12inalcr.pdf>.

⁹ Information provided by the IRS indicates that the IRS had determined the eligibility status of two-thirds of the 12 percent that appear eligible and either denied the claim or sent notices of potential eligibility to the tax filer.

- . 2013. “Earned Income Credit (EIC),” Publication 596 for use in preparing 2013 returns. Washington, DC: Internal Revenue Service. <http://www.irs.gov/pub/irs-prior/p596-2013>.
- . 2014. *Compliance Estimates for the Earned Income Tax Credit Claimed on 2006–2008 Returns*. Publication 5162. Washington, DC: Internal Revenue Service. <http://www.irs.gov/pub/irs-soi/EITCComplianceStudyTY2006–2008.pdf>.
- Klerman, Jacob A., and Caroline Danielson. 2011. “The Transformation of the Supplemental Nutrition Assistance Program.” *Journal of Policy Analysis and Management* 30 (4): 863–88.
- Maag, Elaine M. 2015. “Investing in Work by Reforming the Earned Income Tax Credit.” Urban Institute – Brookings Institution Tax Policy Center, May 20. <http://www.taxpolicycenter.org/UploadedPDF/2000232-investing-in-work-by-reforming-the-eitc.pdf>.
- Maag, Elaine; Michael Pergamit; Devlin Hanson; Caroline Ratcliffe; Sara Edelstein; and Sarah Minton. 2015. “Using Supplemental Nutrition Assistance Program Data in Earned Income Tax Credit Administration: A Case Study of Florida SNAP Data Linked to IRS Tax Return Data.” The Urban Institute, Washington, DC, September.
- Olson, Nina E. 2014. *Written Statement of Nina E. Olson, National Taxpayer Advocate, Hearing on Internal Revenue Service Oversight Before the Subcommittee on Financial Services and General Government Committee on Appropriations, U.S. House of Representatives*. <http://www.taxpayeradvocate.irs.gov/userfiles/file/NTA%20Testimony%20on%20IRS%20Oversight%20before%20Subcommittee%20on%20Financial%20Services%20and%20General%20Government%20Committee%20on%20Appropriations.pdf>.
- Pergamit, Michael R.; Elaine Maag; Devlin Hanson; Caroline Ratcliffe; Sara Edelstein; and Sarah Minton. 2014. “Pilot Project To Assess Validation of EITC Eligibility with State Data: Final Report.” Washington, DC: Urban Institute.
- Plueger, Dean. 2009. “Earned Income Tax Credit Participation Rate for Tax Year 2005.” In *Recent Research on Tax Administration and Compliance: Selected Papers Given at the 2009 IRS Research Conference*, edited by James Dalton and Martha Eller Gangi, 151–95. Washington, DC: Internal Revenue Service.
- Scholz, John Karl. 1994. “The Earned Income Tax Credit: Participation, Compliance, and Antipoverty Effectiveness.” *National Tax Journal* 47 (1): 63–87.
- Tax Policy Center. 2015. “Earned Income Tax Credit Parameters, 1975–2015,” <http://www.taxpolicycenter.org/taxfacts/displayafact.cfm?Docid=36> downloaded on July 2, 2015.
- Trippe, Carole, and Jessica Gillooly. 2010. “Noncash Categorical Eligibility for SNAP: State Policies and the Number and Characteristics of SNAP Households Categorically Eligible Through Those Policies.” Washington, DC: Mathematica Policy Research. http://mathematica-mpr.com/publications/PDFs/nutrition/non-cash_snap.pdf.
- United States Department of Agriculture. 2011. “Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2010.” Washington, DC: Food and Nutrition Service, Office of Research and Analysis. <http://www.fns.usda.gov/sites/default/files/2010Characteristics.pdf>.
- U.S. Department of the Treasury. 2011. *Agency Financial Report Fiscal Year 2011*. Washington DC: U.S. Department of the Treasury. <http://www.treasury.gov/about/budget-performance/annual-performance-plan/Documents/FY%202011%20AFR-Final%20Version.pdf>.
- . 2013. *Agency Financial Report: Fiscal Year 2013*. Washington, DC: U.S. Department of the Treasury. <http://www.treasury.gov/about/budget-performance/annual-performance-plan/Documents/2013%20Department%20of%20the%20Treasury%20AFR%20Report%20v2.pdf>.
- White, James R. 2002. “Earned Income Tax Credit Participation.” GAO-02-290R. Washington, DC: Government Accountability Office.

Acknowledgments

The report on which this paper is based was funded by the U.S. Department of Treasury under contract #GS-23F-8198H. We are grateful to our funders, who make it possible for Urban to advance its mission. It is important to note that funders do not determine our research findings or the insights and recommendations of our experts.

Barbara Wiss served as the project officer for the project this brief was based on and has been extremely helpful and supportive. She and Carlton Maryott, who served as the project officer in the first part of the study, were instrumental in securing the participation of the Florida Department of Children and Families. Kara Leibel of the IRS conducted all of the data work with IRS tax records, matched IRS data to the State analysis files, and created tabulations to support the analysis; we are greatly in her debt. Sisi Zhang and Claudia Sharygin, both formerly with the Urban Institute, contributed at different stages of the project. In Florida, Tammy Lary of the Florida Department of Children and Families coordinated obtaining the SNAP data used in this report; Pat Brown was extremely helpful in helping us understand these data. Frank Sammartino, Leonard Burman (both of TPC), John Wancheck (Center on Budget and Policy Priorities), Deena Ackerman, Alan Plumley, and other participants at an IRS-TPC Research Conference hosted on June 18, 2015, provided thoughtful comments on earlier drafts of this paper. All errors in this report are those of the authors.

IRS Preparer-Level Treatment Tests

Karen Masken, Return Preparer Office, Internal Revenue Service¹

Introduction

In 2010, the Internal Revenue Service (IRS) adopted regulations aimed at establishing standards for paid tax return preparers who prepare individual income tax returns (Form 1040). The objective was to improve voluntary compliance through increased oversight of the paid preparer industry with the goal of reducing errors on tax returns. The Return Preparer Office (RPO) was formed to meet this objective. RPO does not have enforcement authority, but rather focuses primarily on education, outreach, and partnering with the paid preparer community. Specifically, the three primary strategic goals of RPO are:

1. Register and promote a qualified tax professional community;
2. Improve the compliance and accuracy of returns prepared by tax professionals; and
3. Support a stakeholder-focused culture that encourages voluntary compliance and continuous improvement.

One effort to meet these strategic goals was a new requirement, effective January 1, 2011, that all paid preparers who prepare Form 1040 returns: register with the IRS; obtain a preparer tax identification number (PTIN); and enter the PTIN exclusively as the preparer identifying number on the returns they complete. Previously, PTINs had been optional.

Another effort involved a multi-year study that started in 2012 to determine the effect on tax return preparation accuracy of various treatments applied to paid tax return preparers. Because IRS has historically focused on taxpayer-level treatments, there is currently only a limited understanding of how preparer-level treatments affect change in preparer and client tax compliance. The goal of the multi-year study is to understand what treatments are effective on different segments of the noncompliant preparer population in an effort to determine the cost-effectiveness of the treatments. Details on the development and design of the study and results from the first year were presented in an earlier paper.² This paper builds upon the earlier version and includes additional information regarding the second year of the study.

Motivation

IRS resources have become increasingly scarce in the recent budget environment. The driver of the preparer-level treatment approach is that treatment of a single preparer is likely to improve the compliance of many tax returns, increasing the expected return on investment of treatment resources. This is similar to intervening at a wholesale level, rather than at the retail level.

As illustrated in Figure 1, preparer compliance is not a binomial variable, but rather a continuous spectrum. On one end of the spectrum are compliant, well-informed preparers, while on the other end of the spectrum are preparers who willfully perpetrate fraud. In the middle of the spectrum are those who might be unintentionally making errors due to lack of knowledge or those who are willfully noncompliant in their tax preparation but could potentially be moved toward voluntary compliance with a light touch. Moving preparers toward greater voluntary compliance is the most cost-effective action in that it protects revenue by having tax and credits reported correctly on the tax return—rather than IRS trying to recover revenue through examination and collection after the return has been filed, and in many cases, after a refund has been paid.

¹ The views expressed in this paper are those of the author and do not necessarily reflect the official position of the Internal Revenue Service.

² Masken, Karen C. "IRS Preparer-Level Treatment Tests: Results from the First Year of a Multi-Year Study," December 2014.

FIGURE 1. The Preparer Compliance Continuum

Traditionally, IRS has focused its enforcement resources toward the noncompliant and fraudulent end of the spectrum on what are predicted to be the largest, most substantial problems. For less substantial cases and for preparers in the middle of the spectrum, the tradition has been to rely on nontargeted services, primarily broad-based education efforts such as tax forums and webinars. Preparers subject to enforcement are generally subject to examination and in some cases, criminal investigation and prosecution. While these actions are necessary in some circumstances, they are very resource-intensive and costly. Audits of preparers typically involve auditing approximately 30 of their clients (for individual taxpayers, the current audit rate is less than two percent overall). If the audit of the preparer leads to litigation, then the Justice Department becomes involved and a very costly case has to be put together. Typically, the Justice Department handles less than fifty such cases each year. Since there are currently about 700,000 registered preparers, traditional enforcement efforts can reach only a small fraction of them.

Tax return preparers, like taxpayers, have certain rights. While there is a desire by some to take draconian actions against a preparer thought to be noncompliant (“Put them out of business!”, “Shame them publicly!”, etc...), many of these proposed ideas fail to recognize that, in addition to enforcement resource constraints, a preparer is entitled to many of the same rights afforded to individual taxpayers. By law, the IRS cannot share with the general public—or even with the preparer’s clients—the identity of a preparer under suspicion of filing noncompliant returns, as that would violate disclosure statutes and would not afford the preparer due process. The only time a preparer’s identity is disclosed is when a criminal or civil suit is brought against the preparer, at which point the information becomes a matter of public record. Unfortunately, since these are the cases that make headlines, it leaves many with the impression that ALL preparers are nefarious. While there are unequivocally bad actors in the community, RPO’s view is that these are the exception, not the rule, and many noncompliant preparers are in the middle of the compliance continuum.

While preparers toward the noncompliant/fraudulent end of the spectrum may require more expensive and intrusive treatments (e.g., audits/injunctions), finding effective, lower cost treatments for preparers in the middle of the spectrum could have a significant impact on revenue collected. Many preparers in the middle currently go untouched as traditional IRS examination approaches would not be cost-effective and the preparer may not partake in IRS educational services. Many of these preparers simply may not fully understand all the rules and may benefit from education as opposed to enforcement. Others might knowingly make errors believing they go unnoticed, but could be moved toward voluntary compliance by a touch lighter than traditional enforcement. If less expensive treatments are found to be effective, they would allow for a larger number of preparers to be treated and moved toward voluntary compliance by the same finite set of resources, thus protecting more revenue. Therefore, the focus of the treatments being tested is on preparers in the middle of the compliance spectrum. It should also be noted that a preparer is defined as an individual, not the firm or business where the preparer works (e.g., many well-known tax preparation businesses employ many individual preparers).

Issue Selection

The first step in the test development process was to identify which tax compliance issues to address. The following describes the rationale behind the issue choices.

The National Research Program (NRP) is an initiative at IRS that conducts audits on a random sample of individual taxpayers, thereby providing unbiased estimates of compliance for most line items on Form 1040 (individual income tax return). RPO analyzed the NRP results from Tax Years 2006 and 2007 and looked at the tax gap report based on the 2001 NRP study (the most current tax gap report available at the time) to inform the decision as to which issue to address with the treatments in the first year. Not only was the overall

magnitude of the compliance problem considered, but also to what extent returns by paid preparers contributed to noncompliance.

The issue chosen for the initial year of treatment tests was Schedule C net income. RPO found that approximately 75 percent of Schedule C returns in the NRP study were completed by paid preparers, and about 75 percent of those returns had errors. The overall contribution of misreported Schedule C income to the tax gap was approximately \$68 billion in Tax Year 2001, which was almost 30 percent of the individual income tax gap.

In the second year of the study, Schedule C net income continued to be addressed. In addition, RPO decided to address the Additional Child Tax Credit (ACTC), as well. The ACTC is a relatively new tax credit that was introduced in 1998. While RPO estimated from the NRP data that the misreported amount for ACTC was only about 2.5 percent of the Schedule C misreported amount, the ACTC is viewed as a potentially emerging compliance issue. Since approximately 65 percent of returns claiming the ACTC are prepared by paid preparers, it is an issue that may benefit significantly from an educational treatment at the preparer level.

Model Development and Selection Criteria

Schedule C

RPO inherited an established examination plan begun in 2010 that identified preparers for treatment and either subjected them to an Educational Visit or sent them a letter stating there were errors on their returns. The filters used to identify these preparers were developed by subject-matter experts based on their individual experience. In an audit³ by the Treasury Inspector General for Tax Administration (TIGTA), the filters were critiqued for not being data-driven and were identified as a management challenge. TIGTA recommended that a data-driven selection method be developed. RPO agreed with this recommendation, and error detection models using the National Research Program (NRP) data were developed for the Schedule C net income issue.

Because the NRP sampling rate overall is less than 0.01 percent, it is rare for more than one client per preparer to be selected for the random NRP audits. Therefore, the model was initially developed at the taxpayer level and then rolled up to the preparer-level selection criterion. The model was developed using data from the NRP Tax Years 2006 and 2007 studies and tested on the Tax Year 2008 study. In part, this was out of necessity as the Tax Year 2008 data were not yet available when the model was being developed. However, this approach does provide the benefit of assessing how robust the model is to choice of tax year. The model performed as expected on the test data.

After the model was developed, an outside expert performed an independent evaluation of the model and its development. The expert found the methodology was appropriate and that the model was effective. One technique employed to evaluate the model was a confusion matrix wherein the model was compared to a random draw. Based on the confusion matrix, the model was found to be more effective than a random draw by seven standard deviations.

The results from the first year of the study were not available when the second year was implemented. While no refinements to the model for the second year could be made based on new data, some refinements were made by employing the confusion matrix evaluation technique on subsets of the existing data.

To be eligible for the Schedule C test, preparers had to have an active PTIN, had to prepare at least 20 Schedule C returns with at least 15 percent of all their returns containing a Schedule C, and the majority (51 percent or more) of their Schedule C returns had to be flagged by the error-detection model. The requirement that they have at least 20 returns and 15 percent of their total returns with a Schedule C was simply to ensure that Schedule C was prevalent enough in the preparer's business to warrant treatment. Each of a preparer's Schedule C returns was scored using the respective model each year, and those considered high risk were flagged. The Schedule C returns were then aggregated by preparer and the percent of the preparer's Schedule C returns that were flagged was calculated. While it is virtually impossible to solve the endogeneity issue of

³ Implementation of the Return Preparer Visitation Project Was Successful, but Improvements Are Needed To Increase Its Effectiveness, June 29, 2012, TIGTA Audit # 2012-30-068.

whether it is the preparer or the client driving the noncompliance, RPO believed that if the majority of a preparer's Schedule C returns were flagged, then the preparer was likely to be at least partly responsible for the noncompliance. Therefore, if at least 51 percent of a preparer's Schedule C returns were flagged, then the preparer was placed in the treatment pool. There were preparers who met all the criteria except that they did not have an active PTIN. These preparers are both programmatically noncompliant (meaning they did not adhere to the basic requirement that they obtain a PTIN and enter it exclusively as the preparer identifying number on the returns they completed) and are at risk for tax noncompliance as well. RPO believed these preparers needed a different treatment to address both issues, so they were excluded from this set of tests.

ACTC

The initial ACTC analysis was based on information from the Return Transaction File for Calendar Year 2012. At the time the analysis was being performed, there was concern in Congress regarding the validity of children with Individual Tax Identification Numbers (ITINs) being claimed for the ACTC. The IRS issues ITINs to individuals who are required to have a taxpayer identification number but who do not have, and are not eligible to obtain, a Social Security Number. While children with ITINs can be claimed legitimately for the credit, they must meet specific criteria. One interesting phenomenon the data showed was that approximately 1 percent of return preparers accounted for 60 percent of all ACTC ITIN children. In light of this, two separate selection criteria were developed—one for preparers with a significant number of ACTC children with ITINs (dubbed 'ITIN Specialists') and another for preparers who had a significant number of ACTC returns, but not a significant number of ACTC children with ITINs (dubbed 'ACTC Generalists'). In contrast to the Schedule C model, the ACTC model is not an error detection model and does not attempt to make any inferences regarding the compliance of the preparers' returns.

The first set of eligibility criteria for ACTC mirrored those of Schedule C. The preparer had to have an active PTIN, prepare at least 20 ACTC returns, and at least 15 percent of the returns had to have ACTC claims.

To be eligible for the ITIN Specialist treatment, preparers had to have at least 20 returns that had children with ITINs being claimed for ACTC, and at least 15 percent of all their ACTC returns had to include children with ITINs being claimed for ACTC.

For preparers who did not meet the ITIN Specialist criteria, the percent of the preparer's returns that included ACTC was increased from 15 percent to at least 50 percent. This was simply to reduce the pool of eligible preparers and to focus on those with a greater prevalence of ACTC returns.

Treatments

In the initial year of the study, three different types of Schedule C treatments were tested and compared to one another. The first treatment was an "Educational Visit" to the preparers by revenue agents to discuss Schedule C issues found on returns that they had prepared. The second was a "Due Diligence Letter" reminding preparers of their due diligence requirements when preparing returns, and warning that they and their clients might be subject to audit (see Appendix A). The third treatment was a "Continuing Education Letter" with the same message regarding due diligence, but also recommending that, as part of the continuing education required at that time,⁴ the preparer take a minimum of 4 hours of continuing education regarding Schedule C (see Appendix B).

In the second year of the study, two of the three Schedule C treatments were re-employed and an additional two ACTC treatments were tested. For Schedule C, the Educational Visit and Due Diligence Letter tests were repeated. At the time, results from the first year were not in. However, since the first year of the tests, the IRS had lost the *Loving et al.* lawsuit and, as a result, could no longer require continuing education. While the Continuing Education Letter did not specifically state that a preparer *must* take continuing education, RPO

⁴ At the time of the first year of the test, preparers were required to take continuing education. It was later determined by the courts that IRS did not have the statutory authority to impose this requirement (*Loving et al. v. IRS*).

did not want to give the appearance that any type of education was required. Therefore, this treatment was not applied in the second year.

Two different letter treatments were tested for ACTC. In contrast to the Schedule C treatments, the ACTC treatments made no inference regarding the accuracy of the preparers' returns. Instead, they were purely educational and highlighted changes that had recently been made to ACTC filing requirements. Specifically, that what had previously been Schedule 8812 was now Form 8812, *Child Tax Credit*, and required to be filed with the return. One letter, the ITIN Specialist Letter, pointed out changes related to children with ITINs specifically (see Appendix C), while the ACTC Generalist Letter did not (see Appendix D).

Evaluation Method and Test Design

A Randomized Controlled Test (RCT) design was employed for each of the treatment tests, and the effectiveness of each treatment was evaluated using the Difference in Differences (DID) technique. Preparers in each of the treatment pools were randomly assigned to either the treatment or its respective control group. All treatments were applied prior to the start of the respective filing season (2013 the first year, and 2014 the subsequent year) and the determination of the effectiveness of each treatment was based on individual income tax returns filed in the respective years.

For the Schedule C treatments, the determination of effectiveness was based on the percent of clients who were flagged by the error detection model, described previously, in the filing season after the treatment was applied. During the design phase, it was decided that a treatment would be deemed successful for a particular preparer if it lowered the proportion of clients flagged for that preparer by 5 percentage points. While the definition of success will always be somewhat arbitrary, RPO believed it was important to define success during the design phase in order to avoid the perception that success was defined based on preliminary results (and the natural inclination to do so). The overall effectiveness for each treatment was determined by counting the number of successful preparers for each treatment and comparing it to the respective control group.

While preparers were randomly assigned to a treatment or control group, it is important to note that each treatment had a different set of constraints resulting in a different composition of preparers for each treatment and its respective control group. For the Educational Visits, the allocation of resources, namely Revenue Agents, had to be taken into consideration. Rather than drawing a simple random sample from the pool of preparers, a random sample proportionate to the resources available in each of seven IRS-defined geographic areas was drawn. For the Continuing Education Letter, preparers who held a credential (e.g., Certified Public Accountants and attorneys) were exempt from the IRS continuing education requirement in place at the time, as their own credential held them to a higher standard. They were, therefore, excluded from this treatment. The Due Diligence Letter had no constraints. Each of the treatments had its own control group with corresponding constraints; however, the control groups were not mutually exclusive of one another. Finally, each control group incorporated controls for possession of a credential and IRS area.

At the time the treatments were being implemented the first year, Hurricane Sandy struck land, and IRS generally suspends enforcement actions during a natural disaster. The visits were already underway, so preparers in the disaster area States (New Jersey, New York, Rhode Island, and Connecticut) were dropped from the test that year. Neither of the letters had been sent yet, so the test and control groups were redrawn excluding preparers in the disaster area. As a result, the size of the Educational Visit test in the first year is slightly smaller than the two letter tests.

Like the Schedule C tests, preparers in the ACTC tests were randomly assigned to either the test or control group for each treatment (note that, as described previously, the treatment groups are mutually exclusive). Since no error detection models were developed, the success of the tests was based simply on the number of ACTC claims the subsequent year, and for the ITIN Specialist, the number of children claimed for ACTC with ITINs.

Generally speaking, there are five statistical parameters that go into test-size calculations: precision, power, confidence level, the test statistic, and the critical value. For the Schedule C treatment tests specifically, the precision, or the probability of a false-positive (declaring success when it actually failed) was set at the standard

5 percent level. The power of the test, or the probability of a false-negative (declaring failure when it actually succeeded) was set at the standard level of 80 percent. The confidence level was set at 95 percent. The test statistic was the number of successes as defined above. The critical value was 5 percentage points, meaning that there had to be at least a 5 percentage point difference between the test and control to declare the difference statistically significant. The test size, given these parameters, was calculated to be approximately 1,250 preparers for each of the treatment and control groups.

Since the ACTC letters were educational in nature, there was an operational desire to send a larger number than a test would require. Therefore, the test size for each of these letters is larger than what was necessary statistically. There were 3,500 preparers in the ITIN Specialist Letter test and 5,000 in the ACTC Generalist Letter test.

At the initial design phase, an outside expert (different from the one employed to evaluate the model) provided consultation in the development of the test design to ensure the tests would produce the desired information.

Test Preparer Characteristics

Once the pools of preparers were identified, RPO worked with various other IRS offices to remove preparers who had been selected for enforcement efforts (such as criminal investigations). This undertaking is commonly referred to as 'deconfliction' within IRS. The final size of each test pool is presented in Table 1.

For the Schedule C tests, the final pool of preparers after deconfliction in year 1 was 9,600. This dropped in year 2 to 6,800. The main reason for the drop was that in addition to deconflicting with other IRS offices, preparers selected into either the test or control groups in the first year were excluded from the second year of the tests. Previously selected preparers account for 27 percentage points of the 39 percent dropped in year 2.

While the Schedule C tests lost only around 10 percent to deconfliction with other offices, the ACTC tests lost a much larger percentage. The primary reason for this was a simultaneous preparer-level treatment test by the IRS Earned Income Tax Credit (EITC) Program Office. Approximately 80 percent of all returns claiming ACTC also claim EITC, hence the large overlap. While RPO had not initially anticipated dropping so many preparers from the ACTC test pools, it turned out to be beneficial. While the ACTC preparers were not selected based on the risk of the returns they prepared, the preparers selected for EITC treatments were. As explained earlier in the paper, RPO's main focus is the preparers in the middle of the compliance spectrum who would not otherwise be touched. By removing preparers the EITC Program office believed to be high risk, the EITC Program Office helped RPO focus on preparers who are in the middle of the spectrum.

TABLE 1. Final Volume of Preparers Selected into Each Test Pool

Treatment Group	Number of Preparers Before and After Deconfliction		
	Before	After	% Dropped
Schedule C Year 1	10,600	9,600	9%
Schedule C Year 2	11,100	6,800	39%
ACTC ITIN Specialists	5,300	4,000	25%
ACTC Generalists	11,500	5,900	49%

On average, each preparer in the Schedule C pool in the first year prepared approximately 280 Form 1040 returns and 80 Schedule C returns (the medians were 180 and 50, respectively). For the second year, the average number of Form 1040 returns was 250 with an average of 70 Schedule C returns (with medians of 160 and 40, respectively). The average percent of their Schedule C returns flagged by the error detection model was 60 percent for both years.

In comparing the mean number of returns (280 in year 1 and 250 in year 2) and the median number of returns (180 in year 1 and 160 in year 2), it is clear that the volume of returns is highly skewed. The volume ranged from the required minimum of 20 returns to over 5,000 returns in each year. It is important to point

out that, in some instances, a preparer is under the supervision of another preparer who ultimately signs the return. Thus, while all the returns of the supervised preparer will bear that supervisor's PTIN, the supervisor is not necessarily the one who did the actual preparation. It is not possible to tell from the data whether a return was prepared by an unsupervised or supervised preparer.

For the preparers selected into the ACTC ITIN Specialist treatment pool, the average number of returns was 350, and 210 of those, on average, had ACTC claims. On average, 100 of their returns had at least one child with an ITIN being claimed for ACTC, with an average of 230 ACTC children with ITINs in total. The ACTC Generalist pool had an average of 100 returns and 60 ACTC claims. One note of interest is that the preparers who were excluded after deconfliction with the EITC Program Office tended to have a much higher volume of returns. The ITIN Specialists deconflicted had an average of 570 returns with 360 ACTC claims, and the ACTC Generalists had an average of 190 returns with 120 ACTC claims. Again, this illustrates that the final pool of preparers for the ACTC tests reflected preparers who would not normally be touched by IRS enforcement activities because they are too small to warrant enforcement treatment.

Results

This section begins with the general results of the Schedule C tests, followed by results of the ACTC tests. A more detailed analysis of the results for the first year of the Schedule C tests was presented in the earlier paper referred to previously, and is not repeated here.

Schedule C Tests

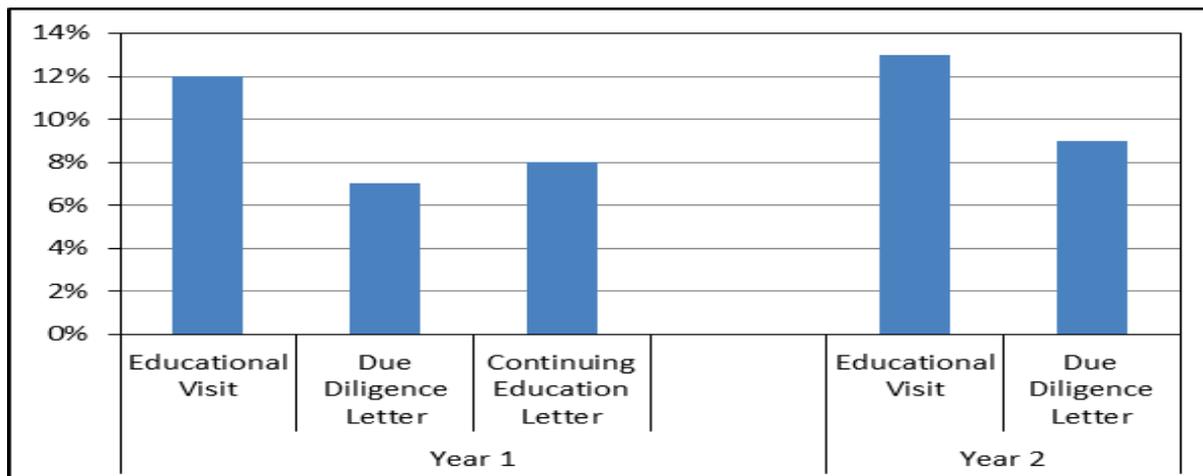
Attrition

One consequence of the treatments could be to effectively remove a preparer altogether from preparing returns. To the extent it put a noncompliant preparer out of business, this would be viewed as a positive result. On the other hand, it could also have an unintended negative consequence of moving the preparer to prepare but not sign returns, making them both programmatically and tax-reporting noncompliant. However, neither of these appears to have occurred to a significant degree.

For the Educational Visit treatment, the attrition rate for the test group in the first year was 3 percent, compared to 2 percent for the control group. However, this difference is not statistically significant. In the second year it was 2 percent for both the test and control groups. The preparers in the Due Diligence Letter test and control groups had a 2-percent attrition rate in both years. The Continuing Education Letter test had a 3-percent attrition rate for both the test and control groups. As previously discussed, the Continuing Education Letter test excluded CPAs and Attorneys, which may explain the slightly higher attrition rate for this group compared to the Due Diligence Letter treatment group since non-CPAs may be less invested in tax preparation.

Success Rate

As explained previously, the tests are evaluated by computing the difference in differences between the test and control groups. The differences, or success rates, for each of the Schedule C treatments are shown in Figure 2. As previously discussed, each of the treatment groups has a different composition and the test size for the Educational Visit treatment was smaller than the other two in the first year. Hence, the counts of success are not directly comparable to one another. While comparing the success rates is more accurate, it is still somewhat problematic in that each treatment had different constraints. However, if the treatments were to become operational, the constraints would remain the same. It is therefore beneficial to RPO to make the comparisons.

FIGURE 2. Success Rates of Schedule C Treatments

The difference in success rates between the test and control groups was highest for the Educational Visit treatment with a 12 percentage point difference between the test and control groups in the first year, and a 13 percentage point difference in the second year. In the first year, the Due Diligence Letter treatment had a success rate of 7 percentage points and the Continuing Education Letter treatment had a success rate of 8 percentage points. While these two are not statistically significantly different from one another, they are significantly lower than the Educational Visit success rate. As noted previously, in the second year the Continuing Education Letter treatment was dropped and the Due Diligence Letter was sent only to credentialed preparers. While the success rate for the Due Diligence Letter treatment in the first year was 7 percentage points overall, when broken down by credential, the noncredentialed preparers had a success rate of only 5 percentage points. Thus, the increase in year 2 to 9 percentage points is likely due to the fact that noncredentialed preparers were excluded from this treatment in the second year. Another explanation for the slightly higher rate for both the Educational Visit and the Due Diligence Letter treatments is that the underlying selection model was improved.

While the Educational Visit treatment was significantly more effective, it is not as cost-effective as either of the letter treatments. It is difficult to get a firm comprehensive cost estimate of a visit. However, the order of magnitude is really all that is necessary. The cost of an Educational Visit is over \$100, while the cost of a letter is under \$1. Thus, the effect of the Educational Visit would need to be much higher than it is in order for it to be cost-effective.

For preparers where the Continuing Education Letter treatment was considered successful, RPO was interested in what percentage took continuing education as a result of the letter. Currently, data available to RPO do not show the topic or the actual date of the continuing education, but there was a 5-percentage point difference in the percent of preparers who took continuing education between the test group (31 percent) and control group (26 percent). Hence, it appears the letter did effectively encourage preparers to take at least some form of continuing education.

Overall, the successful preparers in the first year prepared more than 375,000 individual tax returns with around 100,000 returns including a Schedule C. While a success was defined as a 5-percentage point drop in the number of Schedules C flagged by the error-detection model, the actual average drop across the treatments was 18 percentage points. The estimated number of taxpayers who moved towards voluntary compliance as a result of the three treatments in year 1 is approximately 18,300 taxpayers.

For the second year, successful preparers prepared more than 315,000 individual returns with over 80,000 Schedules C. As in the first year, the average percentage of Schedule C returns flagged by the error detection model dropped by 18 percentage points. The estimated number of taxpayers who moved towards voluntary compliance in the second year was approximately 14,400.

Client Migration

While preparers were the primary focus of the treatments, RPO was also interested in what happened with their clients. As mentioned earlier, there is an issue of endogeneity when trying to determine if the taxpayer or the preparer is driving the noncompliance. To the extent it is the taxpayer driving the noncompliance, then for successfully treated preparers, one might expect these clients to discontinue using that preparer and either go to a new preparer or prepare the return themselves. Clients of successfully treated preparers who were flagged by the error-detection model in both years actually had the same or lower migration rates in the test group than the control group for all three treatments. For the Educational Visit, the test group rate was 22 percent compared to the control group rate of 26 percent. For the Due Diligence Letter, the rate was 27 percent for both the test and control groups. Finally, the rates for the Continuing Education Letter were 30 percent and 33 percent for the test and control groups, respectively. Thus, it appears that the model was effective in identifying preparers who were driving the noncompliance.

Recidivism

While all three treatments were found to be effective for the filing season immediately following the treatment, RPO wanted to know if the effect was persistent in the following year, as well. The recidivism issue was addressed by comparing the success rate 2 years later for preparers who were successful in year 1. In comparing the success rate in year 2 between the test and control groups, there was no statistically significant difference between the test and control for any of the treatments, and therefore recidivism does not appear to be an issue. For the Educational Visit treatment, the success rate for the test was 76 percent, compared to 70 percent for the control. The success rates for the Due Diligence Letter and the Continuing Education Letter test groups were 68 percent and 71 percent, respectively, compared to 70 percent for both control groups.

ACTC Tests

Attrition

Like the Schedule C treatments tests, the ACTC Treatment tests had no effect on the attrition rate for preparers. The attrition rate for the ITIN Specialists was 6 percent for both the test and control groups. The ACTC Generalist preparers were much more volatile with rates of 25 percent for the control group and 26 percent for the test group, but these rates were not statistically significantly different from one another.

Success Rates

As stated previously, both of the ACTC treatments were purely educational in nature and no inference regarding the preparers' compliance was made during the selection process. However, if the treatments effectively move preparers towards voluntary compliance, this can be inferred by simply comparing the change in the number of claims between the test and control groups. Additionally, for the ITIN Specialist Letter treatment a comparison between the test and control group of the number of children with ITINs being claimed provides an indication of the effectiveness of the educational letter.

The average number of ACTC claims for preparers in the ACTC Generalist Letter test group was 4-percentage points lower than the control group. The average total amount of ACTC claimed was also 4-percentage points lower than the control group. The average number of ACTC claims for preparers in the ITIN Specialist Letter test group was also 4-percentage points lower than the respective control group, and the average amount of the ACTC claims was 5-percentage points lower. In addition, the average total number of children with ITINs claimed was 10-percentage points lower than the control group. All of these differences are statistically significant; therefore, all treatments were effective in reducing the number of ACTC claims and moving the preparers towards voluntary compliance.

Preparers in the ACTC Generalist Letter treatment test group prepared more than 480,000 returns in the selection year with more than 285,000 ACTC claims. As a result of the treatment there were an estimated 2,500 fewer ACTC claims. For the ITIN Specialist Letter treatment, preparers had approximately 745,000 ACTC claims with about 800,000 children with ITINs on 1.2 million individual returns. The treatment resulted in an

estimated reduction of 1,800 ACTC claims overall, but a reduction of about 18,700 children with ITINs being claimed on ACTC returns.

Client migration

As with the Schedule C tests, there was some concern that if a preparer told a client that the client was not eligible for ACTC, that client might simply move to a different preparer who would make the claim. In following the clients however, this does not appear to be an issue. For the ACTC Generalist Letter treatment, 48 percent of the clients claiming ACTC the first year stayed with the same preparer, compared to 47 percent of clients for the control group. For the ITIN Specialist Letter treatment, 34 percent of clients claiming ACTC in the first year stayed with the same preparer, compared to 31 percent of the control group.

An additional concern for the ITIN Specialist group was that the child with an ITIN would simply be claimed by someone else on a different tax return. However, the attrition rate for the children with ITINs being claimed by anyone was higher for the test group than the control group (31 percent and 28 percent, respectively), indicating that child migration was not a significant issue.

Summary

Employing a statistically valid Randomized Controlled Test and the Difference in Differences evaluation technique allowed RPO to learn the effectiveness of each of the treatments and make appropriate revisions. The results from the first year of the Schedule C tests indicate all three treatments were effective. The second year indicated that the results were consistent from year to year, and that results from the first year were persistent in the following year. However, while the Educational Visit treatment was more effective than either letter, it is not as cost-effective. As a result of the test, and due to budget constraints within the IRS, the Educational Visit treatment was dropped from the treatment pool in subsequent years. The ACTC tests demonstrated that treatments that are purely educational in nature can also be effective. These educational letters are more cost-effective than the targeted compliance letters since no underlying model needs to be developed to select preparers for treatment. As a result, RPO is currently testing an educational Schedule C treatment in lieu of the target compliance letters, as well as continuing the ACTC letters.

In summary, the initial tests were designed to target specific preparers for compliance treatments. However, the results from the first 2 years of the tests have moved RPO away from targeted treatments and towards more general and less costly educational efforts.

References

- Brown, Robert E., and Mark J. Mazur. (2013) "IRS's Comprehensive Approach to Compliance Measurement," Presentation to the National Tax Association Spring Symposium, Washington, DC.
- Hastie, Trevor; Robert Tibshirani; and Jerome Friedman. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science+Business Media, LLC, New York, NY.
- Hicks, Charles R. (1993) *Fundamental Concepts in the Design of Experiments*. Oxford University Press, New York, NY.
- Internal Revenue Service Data Book*, Fiscal Year 2012. (2013) Publication 55B, Washington, DC.
- Masken, Karen C. (2014) "IRS Preparer-Level Treatment Tests: Results from the First Year of a Multi-Year Study." http://www.irs.gov/pub/irs-utl/irs_preparer_level_treatment_tests.pdf.
- Mazur, Mark J., and Alan H. Plumley. (2007) "Understanding the Tax Gap," Presentation to the National Tax Association Spring Symposium, Washington, DC.
- Treasury Inspector General for Tax Administration. (2012) "Implementation of the Return Preparer Visitation Project Was Successful, but Improvements Are Needed To Increase Its Effectiveness." Report 2012-30-068, June 29, 2012.

Appendix A

Due Diligence Letter



Department of the Treasury
Internal Revenue Service
Return Preparer Office
1122 Town and Country Commons
Chesterfield, MO 63017

Letter	XXXX
Date	Date
To contact us	Phone 1-636-255-1208 8 a.m.- 5 p.m. CT

Tax Return Preparer
Address
City, ST zip

Subject: Reminders about Schedule C preparation

Dear Tax Return Preparer,

A review of tax returns you have prepared in the past year shows that many have a high percentage of traits we believe typically indicate errors in preparing Form 1040, Schedule C, *Profit or Loss from Business (Sole Proprietorship)*. This letter is to remind you of your responsibilities in this area, provide educational assistance, and request that you pay special attention to it next filing season.

Due diligence responsibilities

A paid tax return preparer is expected to take multiple steps to prepare accurate tax returns on behalf of clients. These include reviewing the applicable tax law, and establishing the relevancy and reasonableness of income, credits, expenses, and deductions to be reported on the return. In general, you may rely in good faith without verification upon information furnished by the client. However, you may not ignore the implications of information furnished to, or actually known by you. You must make reasonable inquiries if the information appears to be incorrect, inconsistent with an important fact or another factual assumption, or incomplete.

Schedule C reminders

To prepare accurate Schedules C, you should ask your clients sufficient questions to determine that the expenses claimed are correct and allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they are entitled to claim deductions for nonqualifying expenditures. You should also ask your clients if they have documentation to support the expenses in case receipts are requested by the IRS.

Helpful resources

Based on our analysis of the Schedules C you prepared, we encourage you to review the Schedule C instructions and other IRS publications available at www.irs.gov, keyword: Recommended Reading for Small Businesses. We also encourage you to review Circular 230, Regulations Governing Practice Before the Internal Revenue Service, sections 10.22 and 10.34, titled "Diligence as to accuracy" and "Standards with respect to tax returns and documents, affidavits, and other papers," respectively.

Potential consequences

In the future, both you and your clients may be adversely affected by incorrect returns. Consequences may include any or all of the following:

- If your clients' returns are examined and found to be incorrect, your clients may be liable for additional tax, interest, and penalties.

- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to an unreasonable position can be assessed a penalty of at least \$1,000 per return (IRC section 6694(a)).
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to reckless or intentional disregard of rules or regulations by the tax preparer, can be assessed a penalty of at least \$5,000 per return (IRC section 6694(b)).

We hope this letter has heightened your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Schedules C for your clients.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

Appendix B

Continuing Education Letter



Department of the Treasury
Internal Revenue Service
Return Preparer Office
1122 Town and Country Commons
Chesterfield, MO 63017

Letter	XXXX
Date	Date
To contact us	Phone 1-636-255-1208 8 a.m.- 5 p.m. CT

Tax Return Preparer
Address
City, ST zip

Subject: Recommendation to take CE programs about Schedule C

Dear Tax Return Preparer,

A review of tax returns you have prepared in the past year shows that many have a high percentage of traits we believe typically indicate errors in preparing Form 1040, Schedule C, *Profit or Loss from Business (Sole Proprietorship)*. Therefore you may benefit from continuing education (CE) programs on this topic.

The purpose of this letter is to recommend that as part of your 2013 CE requirement, you take a **minimum of 4 hours** of programs related to business income and/or expenses. Information about CE requirements and a list of IRS approved CE providers is available at www.irs.gov/taxpros/ce.

In addition, as you prepare returns for the next filing season, please pay special attention to your work on Schedule C returns to ensure they are prepared accurately. In particular, make sure you have familiarized yourself with the following areas:

Due diligence responsibilities

A paid tax return preparer is expected to take multiple steps to prepare accurate tax returns on behalf of clients. These include reviewing the applicable tax law, and establishing the relevancy and reasonableness of income, credits, expenses, and deductions to be reported on the return. In general, you may rely in good faith without verification upon information furnished by the client. However, you may not ignore the implications of information furnished to, or actually known by you. You must make reasonable inquiries if the information appears to be incorrect, inconsistent with an important fact or another factual assumption, or incomplete.

Schedule C reminders

To prepare accurate Schedules C, you should ask your clients sufficient questions to determine that the expenses claimed are correct and allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they are entitled to claim deductions for nonqualifying expenditures. You should also ask your clients if they have documentation to support the expenses in case receipts are requested by the IRS.

Helpful resources

Based on our analysis of the Schedules C you prepared, in addition to considering CE programs, we encourage you to review the Schedule C instructions and other IRS publications available at www.irs.gov, keyword: Recommended Reading for Small Businesses. We also encourage you to review Circular 230, Regulations Governing Practice Before the Internal Revenue Service, sections 10.22 and 10.34, titled "Diligence as to accuracy" and "Standards with respect to tax returns and documents, affidavits, and other papers," respectively.

Potential consequences

In the future, both you and your clients may be adversely affected by incorrect returns. We will check whether you complete the continuing education on business income and/or expenses as recommended and we will be looking for improvements in returns you prepare. Incorrect returns may cause any of the following consequences:

- If your clients' returns are examined and found to be incorrect, your clients may be liable for additional tax, interest, additions to tax, and penalties.
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to an unreasonable position can be assessed a penalty of at least \$1,000 per return (IRC section 6694(a)).
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to reckless or intentional disregard of rules or regulations by the tax preparer, can be assessed a penalty of at least \$5,000 per return (IRC section 6694(b)).

We hope this letter and your continuing education focus on Schedule C preparation will heighten your awareness of your responsibilities and help ensure you prepare accurate Schedules C for your clients.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

Appendix C

ACTC ITIN Specialist Letter

Tax Return Preparer
Address
City, ST zip

Dear Tax Return Preparer,

Our information indicates that you prepared income tax returns for Tax Year 2012 claiming the Child Tax Credit and Additional Child Tax Credit. The procedures for claiming these credits were modified for 2012 with the creation of Form 1040, Schedule 8812, *Child Tax Credit*. As a paid preparer, you are responsible for ensuring your clients' returns are accurate.

The purpose of this letter is to provide educational assistance and promote awareness of the new Schedule 8812 requirements. Schedule 8812 has four parts and your clients' circumstances will determine which parts must be completed. We ask that you pay special attention to the requirements for each part, including the requirements for claiming the Child Tax Credit and Additional Child Tax Credit for children with an Individual Taxpayer Identification Number (ITIN) next filing season.

ITIN and Child Tax Credit reminders

To prepare an accurate Schedule 8812, you must ask your clients relevant and probing questions to help you determine if the credit is allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they can claim the credit for ineligible dependents. In general, to be a qualifying child for purposes of the child tax credit, the child must be a citizen, national, or resident of the United States. You should ask your clients questions to identify whether a child with an ITIN meets the substantial presence test for establishing residency, if the child is not a U.S. citizen or U.S. national.

Due diligence responsibilities

A paid tax return preparer must take multiple steps to prepare accurate tax returns on behalf of clients. These steps are a preparer's due diligence and include reviewing the applicable tax law to establish the relevance and reasonableness of income, credits, expenses, and deductions on a return. In general, you can rely in good faith without verification on information your client provides. However, you can't ignore the implication of the information you have. You must make reasonable inquiries if the information appears to be incorrect, inconsistent, or incomplete.

Helpful resources

Specific information about the Schedule 8812 is available on our website at www.irs.gov, keyword: Child Tax Credit. In addition, we recommend you review:

- Schedule 8812 instructions
- Circular 230, Section 10.22, Diligence as to accuracy
- Circular 230, Section 10.34, Standards with respect to tax returns and documents, affidavits, and other papers

I hope this letter has increased your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Child Tax Credit claims for your clients.

Thank you for your attention to this matter.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

Appendix D

ACTC Generalist Letter

Tax Return Preparer
Address
City, ST zip

Dear Tax Return Preparer,

Our information indicates that you prepared income tax returns for Tax Year 2012 claiming the Child Tax Credit and Additional Child Tax Credit. The procedures for claiming these credits were modified for 2012 with the creation of Form 1040, Schedule 8812, *Child Tax Credit*. As a paid preparer, you are responsible for ensuring your clients' returns are accurate.

The purpose of this letter is to provide educational assistance and promote awareness of the new Schedule 8812 requirements. Schedule 8812 has four parts and your clients' circumstances will determine which parts must be completed. We ask that you pay special attention to the requirements for each part next filing season.

Child Tax Credit reminders

To prepare an accurate Schedule 8812, you must ask your clients relevant and probing questions to help you determine if the credit is allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they can claim the credit for ineligible dependents. In general, to be a qualifying child for purposes of the child tax credit, the child must be a citizen, national, or resident of the United States. You should ask your clients questions to identify whether a child meets the substantial presence test for establishing residency, if the child is not a U.S. citizen or U.S. national.

Due diligence responsibilities

A paid tax return preparer must take multiple steps to prepare accurate tax returns on behalf of clients. These steps are a preparer's due diligence and include reviewing the applicable tax law to establish the relevance and reasonableness of income, credits, expenses, and deductions on a return. In general, you can rely in good faith without verification on information your client provides. However, you can't ignore the implication of the information you have. You must make reasonable inquiries if the information appears to be incorrect, inconsistent, or incomplete.

Helpful resources

Specific information about the Schedule 8812 is available on our website at www.irs.gov, keyword: Child Tax Credit. In addition, we recommend you review:

- Schedule 8812 instructions
- Circular 230, Section 10.22, *Diligence as to accuracy*
- Circular 230, Section 10.34, *Standards with respect to tax returns and documents, affidavits, and other papers*

I hope this letter has increased your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Child Tax Credit claims for your clients.

Thank you for your attention to this matter.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

5



Appendix

Conference Program

**An IRS-TPC Research Conference:
Improving Tax Administration Through Research-Driven Efficiencies
Urban Institute, 2100 M Street, N.W., Washington, DC • June 18, 2015**

Program

8:30 – 9:00 Check-in

9:00 – 9:30 Opening

Welcome **Eric Toder** (Co-Director, Tax Policy Center) and
Alain DuBois (Acting Director, IRS Office of Research, Analysis, and Statistics)

Opening Remarks **John Koskinen** (Commissioner, IRS)

9:30 – 11:00 **Session 1: Innovative Methods for Improving Resource Allocation**

Moderator: **Melissa Vigil** (IRS, RAS, Office of Research)

- Estimating Marginal Revenue/Cost Curves for Correspondence Audits
Ron Hodge, Alan Plumley, Kyle Richison, and Getaneh Yismaw (IRS, RAS), and Nicole Mizek, Matt Olson, and H. Sanith Wijesinghe (MITRE Corporation)
- Examining the TDA Collectability Curve: How Does the Aging of TDA Delinquencies Impact Dollars Collected?
Joe Saldana, Jeff Wilson, and Tom Beers (IRS, Taxpayer Advocate Service)
- Analysis of Flow-Through Entities Using Social Network Analysis Techniques
Rahul Tikekar, Ririko Horvath, and Larry May (IRS, RAS), Ashish Agarwal and Shannon Chen (University of Texas at Austin)

Discussant: **Arnie Greenland** (Robert H. Smith School of Business, University of Maryland)

11:00 – 11:15 Break

11:15 – 12:45 **Session 2: Taxpayer Responses to Rules and Enforcement**

Moderator: **Ron Hodge** (IRS, RAS, Office of Research)

- Taxpayer Behavior under Audit Certainty
Benjamin Ayers and Erin Towerly (Univ. of GA), and Jeri Seidman (Univ. of TX at Austin)
- 2011–2012 Schedule M-3 Profiles of Schedule UTP Filers by IRC Section Cited
Charles Boynton, Ellen Legel, and Lisa Rupert (IRS, LB&I), and Portia DeFilippes (Treasury Office of Tax Analysis)
- Individual Nonfilers and IRS-Generated Tax Assessments: Revenue and Compliance Impacts of IRS Substitute Assessment When Taxpayers Don't File
Saurabh Datta, Stacy Orlett, and Alex Turk (IRS, SB/SE)

Discussant: **Danielle Higgins** (Baruch College, City University of New York)

12:45 – 1:30	Keynote Speaker
1:30-3:00	<p>Session 3: Improving Tax Administration by Understanding Taxpayer Behavior</p> <p><u>Moderator:</u> Rahul Tikekar (IRS, RAS, Office of Research)</p> <ul style="list-style-type: none"> • Behavioural Insights To Improve Tax Compliance: Short-Term Impacts from a Randomised Experiment in Guatemala <i>Stewart Kettle (University of Bristol), Marco Hernandez (World Bank), Simon Ruda (Behavioural Insights Team), and Michael Sanders (Harvard University and University of Bristol)</i> • 2014 Taxpayer Choice Model (TCM): Designing Digital Communication Products To Reduce Phone and Mail Inventory <i>Courtney Rasey and Mackenzie Wiley (IRS, W&I)</i> • Examining and Addressing Taxpayer Expectations for Affordable Care Act (ACA) Automated Information <i>Ariel Wooten and Marisa McDaniels (IRS, W&I)</i> <p><u>Discussant:</u> Elaine Maag (The Urban Institute)</p>
3:00 – 3:15	Break
3:15 – 4:45	<p>Session 4: Helping Taxpayers Get It Right</p> <p><u>Moderator:</u> Theresa Pattara (H&R Block)</p> <ul style="list-style-type: none"> • The Nontaxable Combat Pay Election and the Earned Income Tax Credit <i>Suzanne Gleason and Patricia Tong (Treasury Office of Tax Analysis)</i> • Pilot Project to Assess Validation of EITC Eligibility with State Data <i>Michael Pergamit, Elaine Maag, Devlin Hanson, Caroline Ratcliffe, Sara Edelstein, and Sarah Minton (The Urban Institute)</i> • IRS Preparer-Level Treatment Tests <i>Karen Masken (IRS, Return Preparer Office)</i> <p><u>Discussants:</u> Deena Ackerman (Treasury Office of Tax Analysis) and Dayanand Manoli (University of Texas)</p>
4:45 – 5:00	<p>Wrap-up</p> <p>Janice Hedemann (Conference Chair, IRS, RAS, Office of Research)</p>

Please provide input about today's conference by completing one of the evaluation forms or doing it online at:
<http://tpc.io/TPC-IRS-survey>