Papers given at the

7th Annual Joint Research Conference on Tax Administration

Cosponsored by the IRS and the Urban-Brookings Tax Policy Center

Held at the Urban Institute
Washington, DC
June 21, 2017

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

*Prepared under the direction of Barry W. Johnson, Director, Statistics of Income Division

Foreword

This edition of the IRS Research Bulletin (Publication 1500) features selected papers from the IRS-Tax Policy Center (TPC) Research Conference held at the Urban Institute in Washington, DC, on June 21, 2017. Conference presenters and attendees included researchers from many 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 Ben Herndon, the IRS Director of Research, Applied Analytics, and Statistics. The remainder of the conference included sessions on identifying corporation tax avoidance, realizing the potential of tax enforcement, the role of incentives in individual compliance, and creative uses of nontax data. The keynote speaker was Peter Merrill from PricewaterhouseCoopers LLP, who offered his insights on current tax issues.

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

This IRS 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, Alicia Miller, and Rahul Tikekar (RAAS); Terry Ashley (Taxpayer Advocate); John Hribar (Human Capital Office Research), Charles Boynton (Large Business & International Division); Sabyasachi Guharary (RICS); Saima Mehmood and Ariel Wooten (Wage & Investment Division); and Joe Rosenberg (Tax Policy Center). In addition, Blake Greene and Ann Cleven from the Tax Policy Center and Linda Addison from the IRS RAAS Data Management Division oversaw numerous details to ensure that the conference ran smoothly.

This volume was prepared by Paul Bastuscheck, Camille Swick, and Lisa Smith (layout and graphics) and Shannon Cook and Beth Kilss (editors), 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.

Ben Herndon
Director, IRS Office of Research, Applied Analytics, and Statistics

7th Annual IRS-TPC Joint Research Conference on Tax Administration

Contents

4. Creative Use of Nontax Data Sources

5. Appendix

1

Identifying Corporation Tax Avoidance

Lisa De Simone ◆ Lillian F. Mills ◆ Bridget Stomberg

Ted Black ◆ Amy Dunbar ◆ Andrew Duxbury ◆ Thomas Schultz

Paul Demeré ◆ Michael P. Donohoe ◆ Petro Lisowsky

Using IRS Data To Identify Income Shifting to Foreign Affiliates[1]

Lisa De Simone (Stanford Graduate School of Business), Lillian F. Mills (The University of Texas at Austin), and Bridget Stomberg (Indiana University)[2]

We use confidential Internal Revenue Service (IRS) data on the magnitude of U.S.-foreign intercompany transactions to develop a financial statement-based measure of the likelihood that U.S. multinational entities (MNEs) make net intercompany payments out of the U.S. Descriptive analysis shows that although sample firms report net *inbound* intercompany payments on average, high tech firms and small firms report average net outbound payments. The determinants of net outbound payments vary with size, but the likelihood that a firm reports net outbound payments is positively related to high tech operations and income tax incentives across all firms. Supplemental analyses show that firms with net outbound payments historically have not been more likely to be audited or assessed additional taxes upon IRS audit. Our study provides a validated measure based on publicly available data that researchers, investors, and policymakers can use to infer a substantial form of income shifting.

Manuscript available on SSRN at http://ssrn.com/abstract=2477537.

[1] The Internal Revenue Service (IRS) provided confidential tax information to one of the authors pursuant to provisions of the Internal Revenue Code that allow disclosure of information to a contractor to the extent necessary to perform a research contract for the IRS. None of the confidential tax information received from the IRS is disclosed in this treatise. Statistical aggregates are used so that no specific taxpayer can be identified from information supplied by the IRS. All opinions are those of the authors and do not necessarily reflect the views of the IRS.

[2] We thank Kathleen Andries, Ben Ayers, Mary Barth, Phil Berger, John Campbell, Tim Dowd (discussant), Nadine Ebert, Matthew Ege, Robert Holthausen, Ross Jennings, Becky Lester, Petro Lisowsky, Tracie Majors, Kevin Markle, Ed Maydew (editor), Peter Merrill, Tom Neubig, John McInnis, Kathleen Powers, John Robinson, Leslie Robinson, Doug Shackelford (discussant), Jeri Seidman, Joel Slemrod, Jake Thornock, Eric Toder, Robert Ullmann, Luke Watson (discussant), Ryan Wilson (discussant), Brian Williams (discussant), Yong Yu, and workshop participants at the University of Texas at Austin, University of North Carolina, the 2017 IRS-TPC Joint Research Conference, Stanford Accounting Summer Camp 2015, the EIASM 2nd Workshop on Current Research in Taxation, the 2012 AAA Annual Meeting, the 2014 National Tax Association Annual Meeting, the 2012 Oxford Doctoral Symposium, and the 2013 UNC Tax Doctoral Conference for valuable comments and suggestions on prior versions of the paper that did not include IRS data. We also thank Lynn Willden, VP of Tax at Insight Enterprises, for helpful discussions. De Simone acknowledges funding from the Stanford Graduate School of Business. Mills acknowledges funding from the Red McCombs School of Business at the University of Texas at Austin and the Beverly H. and William P. O'Hara Chair in Business.

Income-Shifting by U.S. Multinational Corporations

Amy Dunbar (University of Connecticut), with Ted Black (IRS Research, Applied Analytics and Statistics), Andrew Duxbury (James Madison University), and Thomas Schultz (Western Michigan University)[1]

I. Introduction

We explore the use of new foreign tax information reporting (FTIR) of payments to U.S. multinational corporations to estimate tax compliant income-shifting. The new FTIR data reports payments made by foreign-source payors to U.S. residents. This research is part of a larger effort to estimate the offshore component of the corporation income tax gap. The OECD's (Organization for Economic Cooperation and Development) Base Erosion and Profit Shifting (BEPS) project and "country-by-country" (CbC) reporting recommendations address concerns regarding the ability of multinational enterprises (MNEs) to engage in tax avoidance by exploiting gaps in international tax rules. In particular, global tax authorities find it challenging to fairly and consistently measure the economic value created by a multinational corporation (MNC) in a given country because transfer pricing and cost allocation strategies can artificially shift the recognition of profits to relatively low-tax or no-tax jurisdictions. We suggest that FTIR data provide a third-party source of tax data that will make Internal Revenue Service (IRS) operations more effective by identifying compliant corporate taxpayers.

We first estimate income-shifting by U.S. MNCs to/from their controlled foreign corporations (CFCs), using an N-state income-shifting model that computes after-tax rates of return for each MNC and CFC. We do this by equalizing the rates of return between the MNC and CFCs within the MNC group by adjusting profits, thereby removing the income shift (IS), and then aggregate the IS across all entities within the N jurisdictions each year to provide income-shifting estimates. We use data from the tax returns: the MNC's Form 1120 and the CFC's Form 5471 (*Information Return of U.S. Persons With Respect to Certain Foreign Corporations*) e-filed with the IRS from 2007–2015.[2]

Next, we use FTIR data that include information provided by payors to withholding agents in reporting countries[3] as a source to verify offshore income/deductions reported by U.S. MNCs and their related CFCs. The FTIR data are available from 1999 through 2012. We use data from 2007–2012 to match with e-file data. Thus, although we estimate income-shifting for 2007–2015, we can address compliance only for 2007–2012.

The FTIR data report the payor and recipient if the beneficial owner is a U.S. taxpayer. The data come from the foreign tax administration offices of approximately 30 countries, of which two-thirds designate whether the recipient is a corporation. We identify CFC payors that make payments to a corporate recipient who is also a member of the CFC's MNC group. The result of the payor and recipient matching is an intercompany matched dataset that allows us to compare Federal income tax return data and foreign information return data.

Finally, we develop two compliance measures to support inferences about compliant taxpayers, in contrast to most previous tax compliance research that relies upon evidence about noncompliance. These measures are based on two of the traditional components of the tax gap: filing compliance and reporting compliance. The

[1] We thank our discussant, Tim Dowd, Congressional Joint Committee on Taxation staff, for helpful comments. We also thank Mary-Helen Risler, IRS-RAAS, Charles Christian, Professor Emeritus, Arizona State University, and William Trautman, IRS LB&I, for their help. This paper results from work conducted under IRS contract # TIRNE-11-E-00052. The opinions expressed reflect those of the authors and not of the IRS.

[2] Treasury Regulations § 301.6011-5(a)(1) and (f) provide that a corporation must e-file its Form 1120 if it is required to file at least 250 returns during the calendar year ending with or within its taxable year and it reports "total assets at the end of the corporation's taxable year that equal or exceed $10 million on Schedule L of their Form 1120." The IRS tax return database contains tax return information for electronically filed business tax returns since 2006.

[3] Foreign information return data are provided by the source country to the residence country regarding various categories of income. See http://www.oecd.org/tax.

exploratory results provide global tax authorities with a basis for evaluating income-shifting risks and effectively deploying audit resources.

"Income-shifting," referred to as Base Erosion and Profit Shifting (BEPS)[4] by the OECD, has been studied since at least the early 1970s. These early studies generally documented intrafirm trading patterns that suggested tax motivated income-shifting (Jenkins and Wright (1975), Kopits (1976)). Since that time, studies have generally focused on estimating the amount of income-shifting, its sensitivity to changes in tax rates, and the influence of tax havens.

All studies struggle to determine the true amount of income shifted because "unshifted" profit or loss is unobservable, and therefore they use various methods to approximate unshifted profit. Several studies have attempted to approximate unshifted profit by calculating the amount of profit that should accrue to a firm based on labor, capital, and productivity, the method first used in Hines and Rice (1994). In estimating tax-motivated income-shifting, Hines and Rice (1994) add a tax incentive variable to the production function to estimate its sensitivity to tax rates. A few studies used a variation of this method by examining profit shocks including Dharmapala and Riedel (2013) and Markle (2015).

This stream of research also focuses on calculating the profit elasticity to tax rate changes. Most of these studies followed the main tests used in Hines and Rice (1994) and assumed a linear form. Dowd, Landefeld, and Moore (2017) challenge this linear assumption and find that prior studies overestimated the elasticity at higher tax rates and significantly underestimated elasticity at lower tax rates, suggesting that firms shift additional profits to low-tax jurisdictions when the benefits are at their greatest.

Christian and Schultz (2005) use a different approach based on the theoretical work by Rousslang (1997), which assumes the same marginal after-tax return on assets (ROA) across jurisdictions. Deviations from this norm may be tax-motivated. Using U.S. tax return data from 2001, they estimate that from $58 to $111 billion of income was shifted out of the United States. They identify several limitations in their study, including the use of a two-state model, which combines all foreign jurisdictions into one when comparing foreign and domestic ROAs.

Collins, Kemsley, and Lang (1998) use a two-state model to estimate the effect of the foreign tax rate and foreign tax credit position on foreign sales using publicly available data. Klassen and LaPlante (2012) use the same methodology as Collins, *et al.* (1998), but replace annual measures of tax and profitability with 5-year measures. They estimate firms with foreign tax rates lower than the U.S. rate increased income shifted out of the U.S. by approximately $10 billion over their sample period.

One of the major changes in U.S. tax law during the span of the literature is the advent of the "Check the Box" (CTB) regulations, which allowed MNCs to create a "disregarded" entity (DE) that was recognized as a corporation in the foreign jurisdiction but was disregarded as such by the United States. CTB broadened the ways in which U.S. multinationals could shift income among their foreign subsidiaries by enabling passive income (fixed, determinable, annual, or periodic (FDAP) income) to be paid from high-tax subsidiaries to low-tax subsidiaries while avoiding Subpart F, if structured properly. Several studies have examined the extent to which firms engage in additional income-shifting using DEs. Grubert (2012) uses a CTB indicator variable and finds that it contributed a 1-to-2-percentage-point decline in foreign effective tax rates. Altshuler and Grubert (2005) find similar results and estimate the total foreign tax reduction achieved to be approximately $7 billion by 2002.

We add to this literature in several ways. We first extend the two-state income-shifting analysis presented by Christian and Schultz (C&S) at the 2005 IRS Research Conference to an N-state shifting analysis. We then compare estimates of the income shifted from MNCs to: (1) CFCs in information reporting versus nonreporting countries; (2) CAP-owned CFCs versus non-CAP CFCs in any foreign country; and (3) CFCs that own a DE versus non-DE CFCs.

[4] We use the terms "income" and "profit" interchangeably.

Finally, we create two foreign tax information reporting compliance (FTIRC) measures based on a comparison of FTIR and IRS data. The first identifies the percentage of FTIR recipients relative to matched e-file recipients. The second compares information reporting data with reported income tax return data to reflect the correspondence of what CFCs report to the IRS vis-à-vis what the CFCs report to the tax administration in the source country. We use this second measure to infer compliant offshore income-shifting for the FTIR countries.

II. Income-Shifting Analysis

A. Income-Shifting Framework: N-State Model

Christian and Schultz (2005) posit convergence to a single capital factor price across two jurisdictions. That is, a corporation will allocate capital resources such that the marginal after-tax rate of return on capital investment (ATROC) will be equal between any two jurisdictions.[5] Our current research uses an N-state income-shifting model (ISM) that estimates the income shift variable, IS, to/from the foreign jurisdictions and the United States. Appendix A shows the derivation of the N-state model.

For example, if N = 3, the income-shifting framework posits convergence to a single capital factor price, ATROC, across three jurisdictions. If there is equivalence among all three jurisdictions in ATROC when adjusting for shifting, then:

$$\frac{[Y_d + IS_T] * (1-t_d)}{K_d} = \frac{[Y_1 - IS_1] * (1-t_1)}{K_1} = \frac{[Y_2 - IS_2] * (1-t_2)}{K_2}$$

where:

Y_d = pretax domestic income

Y_1 = pretax foreign jurisdiction 1 income

Y_2 = pretax foreign jurisdiction 2 income

K_d = domestic capital stock

K_1 = foreign jurisdiction 1 capital stock

K_2 = foreign jurisdiction 2 capital stock

t_d = tax rate domestic (Td/Yd) where T = level of taxes

t_1 = tax rate foreign jurisdiction 1

t_2 = tax rate foreign jurisdiction 2

IS_T = total income shift [IST = IS1 + IS2]

IS_1 = income shift jurisdiction 1

IS_2 = income shift jurisdiction 2.

The ISM computes ATROC for each MNC and CFC, then adjusts for the income shift (IS) to equalize the ATROCs, and finally aggregates the IS across all firms within N jurisdictions within a year to provide income-shifting estimates.

We cannot necessarily infer noncompliant income-shifting from the observation of unequal ATROCs. A higher-than-normal CFC ATROC, which reflects profit maximizing behavior in the ISM, corresponds to income being shifted from the U.S. parent to the CFC. However, a higher-than-normal ATROC can also reflect effective supply chain management, political risks, regional macroeconomic conditions, or other factors that raise the variability of returns in a given jurisdiction. Thus, a higher-than-normal ATROC may not be indicative of income-shifting that involves tax-motivated compliance risks.

[5] Rousslang (1997) uses a before-tax rate of return measure. We consider only income taxes when computing ATROC. Consistent with Christian and Schultz (2005), we use average rate of return as a proxy for the unobservable marginal rate of return.

While the ISM does not provide definitive income-shifting compliance inferences, it does provide estimates of the aggregate level of profit shifting offshore. The actual rate of return realized by some firms will be higher or lower than what the model predicts in a given jurisdiction, but on average and in the aggregate, the model's estimate of the total level of income shifted is likely to approximate what is actually realized.

B. MNC/CFC Observations

The dataset used for the ISM is comprised of tax returns filed by U.S. corporations as they appear in the e-file and often includes reporting error and data quality exceptions. CFC information returns, in particular, have numerous data quality exceptions that limit the number of usable CFC records for estimation of the ISM. The approach applied in this analysis is to delete CFC records with severe data quality exceptions. Consequently, a large number of CFC records are excluded from the analysis,[6] but enough records remain to provide reasonable inferences about compliant income-shifting.

1. MNC Industry Distribution

We identify 20,337 MNCs with 859,530 CFCs over 2007–2015. We selected only Forms 5471 that were Category 4 filers: U.S. corporations that have control of a foreign corporation (CFC) at some time during the tax year.[7] The final sample consists of 14,428 MNCs with 359,515 CFCs.

Table 1 shows the industry distribution of the parent company MNC over the sample period (for every other year beginning with 2007), based on the NAICS code on Form 1120. We do not have a balanced panel, so the unique EIN column represents the total number of unique MNCs that filed at least one Form 5471.

The industry distribution is similar for the original sample relative to the final sample. The manufacturing industry includes 26.5 percent of the MNCs, and controls 38.7 percent of the CFCs, whereas the professional services industry includes 24.8 percent of the MNCs, but controls only 12.2 percent of the CFCs.

2. Determination of Income-Shifting Database

Table 2 reports the determination of the income-shifting database, showing the number of unique MNCs, the total CFCs across the years, and CFCs within each year. We delete observations for problems with country identification, capital or income measures, tax rate measures, and rate of return measures. The income-shifting variables are defined in Appendix C. For the ISM, we determine the MNC's capital (K_d), income (Y_d), and taxes (T_d) and the CFC's country-aggregated K_f, Y_f, and T_f.

Country. The ISM solves for the equalizing ATROC between the U.S. parent and its CFC on a country-by-country basis. Thus, we must be able to identify the CFC's country. Each Form 5471 reports the country of incorporation (COI) in Box 1(c) and the principal place of business (PPB) in Box 1(e). We define our country variable CNTRY as PPB, or if PPB is missing, we use COI. If COI is also missing, or CNTRY is "US," the observation is deleted. Table 2 begins with the deletion of 1,782 CFC observations (68 MNCs) for which we could not determine the country and 1,826 CFC observations (24 MNCs) that listed the country as the United States. As a result, 92 MNCs no longer were in the database in any year. We delete only the observation with the exception; thus if an observation is deleted because of a CFC variable, the MNC that has other CFCs will continue to be included in the database.

Capital (K) and Pretax Income (Y). K_d is measured as the sum of buildings and other depreciable assets, depletable assets, land, intangibles, and other assets as reported on Form 1120, Schedule L, Lines 10 through 14, for the domestic jurisdiction. K_f is the sum of Form 5471, Schedule F, Lines 8 through 12, for the foreign jurisdictions.

---

[6] The deletion of CFC records from the ISM database due to data quality exceptions creates a fundamental systemic bias in the income shifting estimates. Deleted CFC records constitute nearly one-half of all positive offshore income (Y_f) of CFCs worldwide.

[7] Control for this purpose means the U.S. corporation owns stock possessing: (1) more than 50 percent of the total combined voting power of all classes of stock of the foreign corporation entitled to vote; or (2) more than 50 percent of the total value of shares of all classes of stock of the foreign corporation. As category 4 filers, the Form 5471 must report intercompany transactions on the Schedule M, *Transactions Between Controlled Foreign Corporation and Shareholders or Other Related Persons*.

TABLE 1. MNC/CFC Industry by Year

Industry	Unique		Year					Total	
	EIN	Percent	2007	2009	2011	2013	2015	5471s	Percent
Agriculture, Forestry, Fishing and Hunting	48	0.3%	52	62	99	92	73	750	0.2%
Mining, Quarrying, and Oil and Gas Extraction	220	1.5%	476	439	463	423	339	3,848	1.1%
Utilities	38	0.3%	113	101	103	68	77	833	0.2%
Construction	105	0.7%	313	314	398	394	375	3,097	0.9%
Manufacturing	3,830	26.5%	15,811	14,439	15,779	15,860	15,062	139,242	38.7%
Wholesale Trade	1,218	8.4%	1,952	2,211	2,768	2,628	2,863	22,515	6.3%
Retail Trade	377	2.6%	562	625	695	800	830	6,194	1.7%
Transportation and Warehousing	206	1.4%	505	503	537	609	639	4,992	1.4%
Information	1,204	8.3%	1,842	2,030	2,340	2,534	2,694	20,895	5.8%
Finance and Insurance	424	2.9%	769	717	837	1,188	1,266	9,062	2.5%
Real Estate and Rental and Leasing	155	1.1%	642	684	689	584	584	5,705	1.6%
Professional, Scientific, and Technical Services	3,573	24.8%	4,151	4,437	4,759	5,255	5,768	43,758	12.2%
Management of Companies and Enterprises	2,343	16.2%	8,354	7,977	9,267	10,326	10,893	83,906	23.3%
Administrative and Support and Waste Mgmt	219	1.5%	553	428	750	840	977	6,368	1.8%
Educational Services	55	0.4%	99	132	158	139	110	1,147	0.3%
Health Care and Social Assistance	115	0.8%	103	95	106	75	98	850	0.2%
Arts, Entertainment, and Recreation	71	0.5%	208	233	208	224	245	2,036	0.6%
Accommodation and Food Services	68	0.5%	158	208	177	176	198	1,598	0.4%
Other Services (except Public Administration)	107	0.7%	123	201	282	304	157	1,796	0.5%
Other	52	0.4%	383	164	20	44	39	923	0.3%
Total	**14,428**		**37,169**	**36,000**	**40,435**	**42,563**	**43,287**	**359,515**	

NOTES: Industry based on NAICS (SCHK_BUS_ACTY_CD). See https://www.census.gov/eos/www/naics/.
Agriculture, Forestry, Fishing and Hunting: 111100–123456;
Mining, Quarrying, and Oil and Gas Extraction: 211110–213112; 21230–22100;
Utilities: 221000–221300;
Construction 233110–308700;
Manufacturing: 33400- 3390, 311110–412278;
Wholesale Trade: 421100–439100;
Retail Trade: 45411. 441110–454390, 455110;
Transportation and Warehousing: 481000–493120;
Information: 51112–51900, 511110–519190;
Finance and Insurance: 52312–52590, 521112–529990;
Real Estate and Rental and Leasing: 531100–537900;
Professional, Scientific, and Technical Services: 54119–54519, 541110–547100;
Management of Companies and Enterprises: 551110–55112;
Administrative and Support and Waste Management and Remediation Services: 561100–562211, 562910;
Educational Services: 611000-611710;
Health Care and Social Assistance: 621111–651112;
Arts, Entertainment, and Recreation: 711100–713910;
Accommodation and Food Services: 72215, 721110–722515;
Other Services (except Public Administration: 81121, 811110–818210;
Public Administration: ≥ 920000.

Values for K_d and K_f are required to be greater than zero for ATROC to be defined. Thus, we delete observations when $K_d <= 0$ or $K_f <= 0$. Most of these deletions result from zero assets in the MNC or CFC; thus the entity may be dormant.

Y_d is measured as taxable income reported on Form 1120, Line 30, less foreign-source dividends reported on Form 1120, Schedule C, lines 13-15. Foreign pretax income, Y_f, is measured as the current earnings and profits reported on Forms 5471, Schedule H, Line 5(d), plus foreign tax reported on Forms 5471, Schedule E, Line 8(d).

Because the effective tax rate is not defined when $Y = 0$, we delete observations where Y_d or Y_f is 0. We delete 434,129 CFC observations (855,922–421,793) and 5,636 MNCs (20,245–14,609).

Tax (T) and Effective Tax Rate (ETR). We next delete observations related to tax expense. The tax variable, T, represents the tax liability associated with an entity's taxable income reported in an individual jurisdiction. The value of T impacts the income-shifting estimates because ATROC is an after-tax concept.

The domestic tax variable, T_d, is equal to the sum of total tax reported on Form 1120, line 31, plus the foreign tax credit reported on Form 1120, Schedule J, Line 6(a), multiplied by the ratio of domestic pretax income, Y_d, to taxable income reported on Form 1120, Line 30.[8] The foreign tax variable, T_f, is defined as the income, war profits and excess profits taxes paid or accrued as reported on Forms 5471, Schedule E, Line 8(d). We delete observations where T<0. Because the returns are original, not amended, returns; we cannot explain negative tax expense.

TABLE 2. Determination of Income-Shifting Dataset: MNCs/CFCs, 2007–2015

Item	Total MNCs	Total CFCs	2007	2008	2009	2010	2011	2012	2013	2014	2015
MNCs	**20,337**		**7,053**	**7,575**	**8,151**	**9,119**	**9,872**	**10,426**	**11,103**	**11,972**	**12,571**
CFCs		**859,530**	**89,381**	**92,188**	**91,482**	**92,734**	**96,304**	**97,366**	**98,104**	**99,846**	**102,125**
Dropped obs.											
Missing CNTRY	68	1,782	355	266	217	185	169	141	138	142	169
CNTRY="US"	24	1,826	305	254	196	185	197	203	171	162	153
Remaining	**20,245**	**855,922**	**88,721**	**91,668**	**91,069**	**92,364**	**95,938**	**97,022**	**97,795**	**99,542**	**101,803**
$K_d = 0$	1,311	60,599	8,635	8,647	7,539	5,542	6,013	6,252	6,109	6,022	5,840
$K_d < 0$	38	4,720	617	854	676	350	615	482	573	269	284
$K_f = 0$	3,699	346,568	35,128	36,840	37,432	39,085	39,079	39,157	38,829	40,053	40,965
$K_f < 0$	19	4,398	451	432	361	422	506	481	514	562	669
$Y_d = 0$	493	13,635	1,015	971	1,012	1,794	1,739	1,861	1,676	1,739	1,828
$Y_f = 0$	76	4,209	491	495	412	436	499	461	478	482	455
Remaining	**14,609**	**421,793**	**42,384**	**43,429**	**43,637**	**44,735**	**47,487**	**48,328**	**49,616**	**50,415**	**51,762**
$t_d < 0$	55	29,133	2,002	2,606	3,899	3,400	3,534	2,954	3,054	3,412	4,272
$t_f < 0$	79	16,770	1,590	1,715	2,074	1,856	1,815	1,826	1,976	1,879	2,039
$tr_d > 1$	3	291	61	15	46	22	9	28	62	22	26
$tr_f > 1$	23	10,405	993	1,032	1,035	1,063	1,137	1,217	1,275	1,290	1,363
tr_d, $tr_f < -1$	21	5,667	569	587	571	577	557	663	686	682	775
Remaining	**14,428**	**359,527**[1]	**37,169**	**37,474**	**36,012**	**37,817**	**40,435**	**41,640**	**42,563**	**43,130**	**43,287**
MNCs by year			4,820	5,079	5,271	5,718	6,195	6,492	6,900	7,336	7,628
$atROA_d > 1$	205	3,724	419	583	504	412	323	308	413	398	364
$atROA_f > 1$	1,075	89,053	9,355	9,620	8,256	9,317	10,104	10,138	10,465	10,833	10,965
$atROA_d < -1$	1,623	9,197	915	906	1,022	828	838	1,016	1,067	1,259	1,346
$atROA_f < -1$	1,432	39,454	3,722	4,414	4,142	3,890	4,302	4,497	4,640	4,875	4,972
Remaining	**10,093**	**218,099**	**22,758**	**21,951**	**22,088**	**23,370**	**24,868**	**25,681**	**25,978**	**25,765**	**25,640**
MNCs by year			3,500	3,550	3,742	4,033	4,291	4,391	4,541	4,679	4,766

[1]The sample is further reduced to 359,515 CFCs by a single MNC's 12 CFCs that produced outlier income-shifting estimates despite truncating ATROC.

The domestic effective tax rate is then computed as domestic tax divided by domestic pretax income, or $ETR_d = T_d / Y_d$. The foreign tax rate is computed as foreign tax divided by foreign pretax income, or $ETR_f = T_f / Y_f$. We also delete observations where computed ETR is greater than 100 percent or less than -100 percent. After

[8] Total tax reported on Form 1120, Line 30, is net of the foreign tax credit (see Form 1120, Schedule J, Line 10), so the foreign tax credit is added back to Total Tax in calculating the domestic tax, T_d.

these deletions, the ISM database includes 14,428 MNCs and 359,527 CFCs. The sample is further reduced to 359,515 by a single MNC's 12 CFCs that produced outlier income-shifting estimates despite adjusting ATROC as described below.

After-Tax Return on Capital (ATROC). The ATROC ratio for an MNC in the U.S. or a foreign jurisdiction is defined as (Y-T)/K. We truncate observations where the ATROC is greater than 100 percent or less than -100 percent. An extremely large shifting estimate results when the foreign ATROC is high because of high income in the numerator and low assets in the denominator. The high foreign ATROC may be the result of an investment in a foreign country that has been expensed for tax purposes, but the investment is yielding large returns. If we dropped observations corresponding to ATROCs outside the range, the sample would be reduced to 10,093 MNCs and 218,099 CFCs. We choose to truncate rather than drop to maintain the sample size.

3. Descriptive Statistics

Table 3 reports the mean and median of the variables for each year. As stated before, we do not have a balanced panel of MNCs, even before the deletion process. The final database has 14,428 MNCs, but the number of MNCs ranges from 4,820 in 2007 to 7,628 in 2015. The number of CFCs ranges from 37,169 in 2007 to 43,287 in 2015. The CFC variables are reported both as aggregates across CFCs at the MNC level and at the CFC level. For example, in 2007, 4,820 MNCs invested on average $349 million across all their foreign CFCs, for a total of $1.68 billion. Each CFC has an average of $45 million in capital, for a total of $1.68 billion (37,169 CFCs * $45M).

After-Tax Return on Capital (ATROC). The ATROC ratio for an MNC in the U.S. or a foreign jurisdiction is defined as (Y–T)/K. We truncate observations where the ATROC is greater than 100 percent or less than -100 percent. An extremely large shifting estimate results when the foreign ATROC is high because of high income in the numerator and low assets in the denominator. The high foreign ATROC may be the result of an investment in a foreign country that has been expensed for tax purposes, but the investment is yielding large returns. If we dropped observations corresponding to ATROCs outside the range, the sample would be reduced to 10,093 MNCs and 218,099 CFCs. We choose to truncate rather than drop to maintain the sample size.

Capital Stock (K). For all MNCs, the mean (median) of K_d is $1,653 ($68) million in 2007 and $1,734 ($36) million in 2015. The values of K_f aggregated at the MNC level represent the offshore investment across CFCs for an MNC, decreasing from a mean (median) of $349 ($5) in 2007 to $286 ($2) in 2015. At the CFC level, the mean (median) is $45 ($1) in 2007 and $50 ($1) in 2015.

Pretax Income (Y). For all MNCs, the mean (median) of Y_d is $96 ($0) million in 2007 and $76 ($0) in 2015. The zero medians indicate that losses are reported in the domestic jurisdiction by one-half of the MNCs. The values of Y_f aggregated at the MNC level represent the offshore pretax income across CFCs for an MNC, decreasing from a mean (median) of $47 ($1) in 2007 to $69 ($0) in 2015. For all CFCs, the mean (median) of Y_f is $6 ($0) million in 2007 and $12 ($0) in 2015. The zero median indicates that losses are also reported on average by the CFCs.

Tax (T) and Effective Tax Rate (ETR). For all MNCs, the mean (median) values for T_d is $37 ($0) with a corresponding ETR of 0.11 (0.00) in 2007. The values in 2015 are $30 ($0) with an ETR of 0.13 (0.00). For all CFCs, the mean (median) value of T_f is $2 ($0) in 2007 with a corresponding ETR of 0.14 (0.07). The values in 2015 are $2 ($0) with an ETR of 0.12 (0.01).

After-Tax Return on Capital (ATROC). The ATROC ratio for an MNC in the United States and a CFC in a foreign jurisdiction is defined as (Y–T)/K. Table 3 reports the untruncated results. The mean (median) value of $ATROC_d$ before truncation is -15,906 percent (1 percent) and after truncation is -5 percent (1 percent) (not tabulated). For CFCs, the mean (median) value of $ATROC_f$ before truncation is 36,921 percent (18 percent) and after truncation is 24 percent (18 percent). In 2015, the mean (median) value of $ATROC_d$ before truncation is -78,451 percent (-1 percent) and after truncation is -16 percent (-1 percent). For CFCs, the mean (median) value of $ATROC_f$ before truncation is 812,374 percent (15 percent) and after truncation is 20 percent (15 percent).

TABLE 3. Descriptive Statistics for Input Variables and Key Ratios: Domestic vs. Foreign

[$ amounts in millions]

Variable	2007			2008			2009		
	N	Mean	Median	N	Mean	Median	N	Mean	Median
K_d	4,820	1,652.86	67.72	5,079	1,559.20	61.71	5,270	1,492.62	51.58
K_f-MNC	4,820	348.86	5.05	5,079	279.91	4.15	5,270	312.81	3.79
K_f-CFC	37,169	45.24	1.35	37,474	37.94	1.25	36,000	45.79	1.29
Y_d	4,820	96.22	0.31	5,079	61.49	-0.20	5,270	35.64	-0.38
Y_f-MNC	4,820	46.84	0.83	5,079	0.60	0.53	5,270	56.32	0.32
Y_f-CFC	37,169	6.07	0.39	37,474	0.08	0.34	36,000	8.25	0.24
T_d	4,820	36.76	0.00	5,079	27.65	0.00	5,270	20.71	0.00
T_f-MNC	4,820	16.06	0.18	5,079	17.91	0.13	5,270	10.74	0.07
T_f-CFC	37,169	2.08	0.03	37,474	2.43	0.03	36,000	1.57	0.02
ETR_d	4,820	0.11	0.00	5,079	0.10	0.00	5,270	0.09	0.00
ETR_f	37,169	0.14	0.03	37,474	0.13	0.02	36,000	0.12	0.00
$atROC_d$	4,820	-159.06	0.01	5,079	-18.24	-0.01	5,270	-8.47	-0.02
$atROC_f$	37,169	369.21	0.18	37,474	12,912.55	0.18	36,000	-1,179.36	0.13

Variable	2010			2011			2012		
	N	Mean	Median	N	Mean	Median	N	Mean	Median
K_d	5,718	1,683.84	58.01	6,195	1,562.84	56.98	6,492	1,638.74	49.37
K_f-MNC	5,718	312.95	3.23	6,195	300.96	3.11	6,492	309.63	2.57
K_f-CFC	37,817	47.32	1.32	40,435	46.11	1.30	41,640	48.27	1.32
Y_d	5,718	64.47	-0.05	6,195	57.05	-0.05	6,492	70.73	-0.01
Y_f-MNC	5,718	99.87	0.42	6,195	80.24	0.34	6,492	80.92	0.27
Y_f-CFC	37,817	15.10	0.33	40,435	12.29	0.32	41,640	12.62	0.30
T_d	5,718	27.97	0.00	6,195	26.64	0.00	6,492	29.69	0.00
T_f-MNC	5,718	14.40	0.08	6,195	15.24	0.07	6,492	13.33	0.06
T_f-CFC	37,817	2.18	0.03	40,435	2.33	0.03	41,640	2.08	0.03
ETR_d	5,718	0.14	0.00	6,195	0.14	0.00	6,492	0.15	0.00
ETR_f	37,817	0.13	0.02	40,435	0.13	0.02	41,640	0.13	0.03
$atROC_d$	5,718	-3.46	0.00	6,195	-109.58	0.00	6,492	-10.90	0.00
$atROC_f$	37,817	27,250.51	0.16	40,435	1,776.09	0.16	41,640	294.56	0.15

Variable	2013			2014			2015		
	N	Mean	Median	N	Mean	Median	N	Mean	Median
K_d	6,900	1,612.36	44.99	7,336	1,693.10	39.83	7,628	1,734.38	36.27
K_f-MNC	6,900	302.18	2.20	7,336	287.44	1.81	7,628	285.50	1.60
K_f-CFC	42,563	48.99	1.31	43,130	48.89	1.22	43,287	50.31	1.13
Y_d	6,900	73.43	-0.04	7,336	86.68	-0.08	7,628	75.57	-0.16
Y_f-MNC	6,900	70.85	0.23	7,336	84.71	0.18	7,628	68.86	0.13
Y_f-CFC	42,563	11.49	0.29	43,130	14.41	0.28	43,287	12.13	0.24
T_d	6,900	29.95	0.00	7,336	33.61	0.00	7,628	30.42	0.00
T_f-MNC	6,900	12.16	0.05	7,336	31.54	0.04	7,628	9.24	0.03
T_f-CFC	42,563	1.97	0.03	43,130	5.36	0.03	43,287	1.63	0.03
ETR_d	6,900	0.15	0.00	7,336	0.14	0.00	7,628	0.13	0.00
ETR_f	42,563	0.13	0.02	43,130	0.13	0.02	43,287	0.12	0.01
$atROC_d$	6,900	-14.56	0.00	7,336	-481.44	-0.01	7,628	-784.51	-0.01
$atROC_f$	42,563	4,836.65	0.15	43,130	1,324.64	0.15	43,287	8,123.74	0.15

C. *Income-Shifting Results*

We estimate income-shifting by MNC by country. We then compute the aggregate income-shifting estimates, IS, for each tax year (see Table 4). IS is an estimated $176 billion in 2007 and rises to $274 billion in 2015. The current specification of the N-state ISM does not estimate the amount of cross-subsidiary shifting between countries. Rather, it attributes all income-shifting to transactions directly between the United States and each jurisdiction.

We examine the income-shifting results by segregating the final database by three attributes: (1) whether the CFC is operating in a reporting or nonreporting country; (2) whether the CFC is controlled by an MNC that has a DE in its corporate structure; and (3) whether the CFC is controlled by an MNC that is a CAP taxpayer.

1. Reporting vs Nonreporting Country Analysis

We report income-shifting measures grouped by CFCs operating in reporting and nonreporting countries.[9] Table 4, Panel A, shows the number of CFCs operating in a reporting versus nonreporting country in each year and their corresponding IS, computed at the MNC level as the aggregate of outbound shifting to its CFCs in reporting versus nonreporting countries. Approximately two-thirds of the CFCs operate in a reporting country (227,853/359,515), but account for only 18 percent of the IS ($368,919/$1,995,737).

TABLE 4. Income-Shifting (IS) Estimates

[$ amounts in millions]

Panel A: IS by CFCs Operating in Reporting/Nonreporting Countries

Year	All CFCs		Reporting		Not Reporting		Not Identified	
	N	Total IS	N	IS	N	IS	N	IS
2007	37,169	$176,355	23,790	$24,878	13,107	$146,910	272	$4,567
2008	37,474	202,984	24,016	33,569	13,258	161,201	200	8,213
2009	36,000	171,306	23,011	60,603	12,871	106,227	118	4,475
2010	37,817	252,807	24,149	53,090	13,580	194,397	88	5,320
2011	40,435	255,712	25,742	66,457	14,576	183,168	117	6,086
2012	41,640	210,140	26,311	17,925	15,210	185,485	119	6,731
2013	42,563	220,053	26,658	19,587	15,784	194,359	121	6,107
2014	43,130	232,109	26,926	45,787	16,080	183,461	124	2,861
2015	43,287	274,271	27,250	47,021	15,916	226,963	121	288
Total	**359,515**	**$1,995,737**	**227,853**	**$368,919**	**130,382**	**$1,582,171**	**1,280**	**$44,648**

MNCs can have CFCs in both reporting and nonreporting countries. The ISM determines IS country-by-country, summing the variables for multiple CFCs in a single country. When the Form 5471 provides an unrecognized country code (possible data error), we treat the country as "Not Identified." N is the number of CFCs that remain in the final income-shifting dataset in each year (see Table 2).

Panel B: IS by CFCs With and Without a Disregarded Entity

Year	All CFCs		CFCs with DEs		CFC no DEs		Not Identified	
	N	IS	N	IS	N	IS	N	IS
2007	37,169	176,355	1,790	110,669	35,025	65,419	354	268
2008	37,474	202,984	1,885	141,030	35,192	61,618	397	337
2009	36,000	171,306	1,890	117,232	33,729	53,875	381	198
2010	37,817	252,807	2,157	194,604	35,647	58,212	13	(9)
2011	40,435	255,712	2,324	178,800	38,102	76,915	9	(3)
2012	41,640	210,140	2,566	174,186	39,071	35,952	3	3
2013	42,563	220,053	2,744	187,534	39,816	32,518	3	1
2014	43,130	232,109	2,821	177,018	40,308	55,652	1	(561)
2015	43,287	274,271	3,000	217,515	40,286	56,705	1	51
Total	**359,515**	**1,995,737**	**21,177**	**1,498,588**	**337,176**	**496,865**	**1,162**	**284**

[9] When Form 5471 provides an unrecognized country code (possible data error), we treat the country as "Not Identified."

2. DE/Non-DE Analysis

A relatively easy way to reduce foreign taxes is to shift income from a high-tax foreign entity to a sister low-tax foreign entity through interest or royalty payments. However, this approach was generally ineffective before the "Check the Box" (CTB) regulations became effective in 1997. Because the income earned by the low-tax foreign entity was considered Subpart F income and immediately taxed in the United States, worldwide tax savings were not realized.[10] The strategic use of CTB can eliminate the recognition of Subpart F income on such transactions. By "checking the box" on related tax forms, the two entities are treated as part of the foreign parent CFC for U.S. tax purposes. Thus the *intercompany* payments become *intracompany* transactions, which typically are not subject to Subpart F.

This strategy takes advantage of a mismatch of entity classification between the U.S. and the foreign jurisdictions. The OECD in its "Action Plan on Base Erosion and Profit Shifting" specifically references "country rules that allow taxpayers to choose the tax treatment of certain domestic and foreign entities could facilitate hybrid mismatches" as a way to achieve unintended tax results.[11] Because CTB can reduce the foreign tax rate, it also increases the incentive for MNCs to shift the recognition of what would otherwise be U.S. income to foreign locations.

We assign DE status to the MNC if any related CFC checks the Form 5471, Schedule G box for the question: "During the tax year, did the foreign corporation own any foreign entities that were disregarded as entities separate from their owners under Regulations sections 301.7701-2 and 301.7701-3."

Table 4, Panel B, shows the number of CFCs with at least one disregarded entity and their corresponding IS. Only 6 percent (21,177/359,515) of the CFCs have at least one disregarded entity, but they have 75 percent ($1,498,588/$1,995,737) of the IS.

3. CAP/Non-CAP Analysis

The Compliance Assurance Process (CAP) is an IRS initiative designed to resolve uncertain tax positions in real-time. This research examines whether 221 firms selected for the CAP program have lower shifting measures compared with a sample of 280 non-CAP firms matched on income, assets, and debt equity ratio, using tax return data.

As of the end of 2015, there were 221 firms that have participated in CAP, which started in 2005 with 16 firms. Of the 221 firms, 36 have left the CAP program. Beginning in 2012, the CAP program moved from an invitation-only program to an application program. Initially, taxpayers were invited by the IRS to apply to the CAP pilot program. The IRS required applicants to have assets greater than $10 million, to have audited financial statements, not to be in litigation with the IRS or any federal or state agency, and to display a general willingness to be transparent and cooperative with the IRS.

Table 5, Panel A, reports the aggregate income-shifting estimates, the sums of IS for each tax year for CAP and non-CAP MNCs. Although we have a matched sample of CAP and non-CAP taxpayers, the Table 5, Panel A, results cannot tell us if CAP MNCs shift less income without normalization of the IS measure.

To be able to determine whether the CAP attribute is correlated with IS, we determine the percentage of income shifted for CAP and non-CAP MNCs. We retain the CAP/non-CAP observations where $Y_d > 0$. We normalize IS by year by dividing by pretax income + income shift (Y_d + IS) for each year. For example, if for one year, cumulative Y_d = $500 and IS = $300, then normalized NIS = 300/(500 + 300) = 0.375.[12] Table 5, Panel B contains the results. For example, in 2007, CAP firms shifted 14 percent of income relative to 21 percent by non-CAP firms. CAP MNCs consistently shift less income on a percentage basis than non-CAP MNCs, suggesting firms that participate in CAP appear to be more compliant taxpayers.

[10] Prior to CTB, firms often minimized foreign taxes by using strategies similar to "Commissionaire" structures that limit profits in high-tax jurisdictions, leaving excess profits in low-tax jurisdictions. While there are other legal aspects to these structures, and other structures firms used, they typically earned a small profit. These structures often had undesired business consequences such as VAT.

[11] OECD (2013).

[12] Alternatively, we could have normalized at the MNC level, but adding percents is problematic.

TABLE 5. Normalized IS from CFCs Controlled by a CAP-MNC vs. Non-CAP-MNCs
[$ amounts in millions]

Panel A: IS to CFCs Controlled by a CAP-MNC or Non-CAP-MNC

Year	CAP		Non-CAP	
	N	IS	N	IS
2007	2,794	25,416	5,386	40,563
2008	2,343	20,507	5,121	44,633
2009	2,695	21,772	4,929	40,430
2010	2,781	30,347	5,329	71,558
2011	2,488	24,947	5,193	70,636
2012	2,583	13,440	5,300	53,470
2013	2,666	23,432	5,346	59,491
2014	2,576	15,973	4,922	57,268
2015	2,576	41,779	4,679	48,865
Total	**23,502**	**217,611**	**46,205**	**486,912**

NOTE: 221 CAP firms were matched with 280 non-CAP firms using income, assets, and debt equity ratio from tax return data.

Panel B: Normalized IS to CFCs Controlled by a CAP-MNC or Non-CAP-MNC

Year	CAP					Non-CAP				
	N					N				
	MNCs	CFCs	IS	Yd	NIS	MNCs	CFCs	IS	Yd	NIS
2007	92	1,529	19,576	118,785	0.141	126	2,923	32,471	119,056	0.214
2008	76	998	12,604	81,065	0.135	115	2,544	37,409	103,731	0.265
2009	73	1,084	15,696	78,268	0.167	101	2,278	25,413	84,750	0.231
2010	88	1,356	24,473	106,205	0.187	109	2,716	61,397	106,321	0.366
2011	83	1,248	21,468	80,342	0.211	111	2,884	64,719	103,570	0.385
2012	89	1,388	14,887	101,455	0.128	109	3,049	46,879	110,736	0.297
2013	92	1,510	20,733	122,747	0.145	109	3,124	53,609	126,083	0.298
2014	94	1,468	13,899	139,405	0.091	103	2,831	50,848	149,238	0.254
2015	94	1,395	35,731	131,757	0.213	93	2,630	40,942	132,409	0.236
Total	**781**	**11,976**	**179,066**	**960,028**		**976**	**24,979**	**413,685**	**1,035,894**	

NOTE: We retain the CAP/non-CAP observations where Y_d >0. We normalize IS at the MNC level by dividing by pretax income + income shift (Y_d + IS) for each MNC. For example, if for one MNC in 2007, Y_d = $500 and IS = $300, normalized NISp = 300/(500 + 300) = .375. To avoid adding percents, we sum Y_d within year for CAPs and non-CAPs. In 2007, NIS = 19,576/(118,785 + 19,576) = .141. With the exception of 2007, CAPs consistently shift less income on a percentage basis than non-CAPs.

III. Foreign Tax Information Reporting (FTIR) Data Analysis

A. FTIR Records

This section introduces the use of FTIR data for corporate tax compliance research and explores how this data can be used to draw tax compliance inferences about income-shifting by CFCs. The information shared with the United States is collected by the host countries as part of their regular tax administration processes.

IRS research has demonstrated that tax compliance among individuals improves when they are subject to accurate information reporting by third parties. This study examines whether corporate taxpayers accurately report information to the IRS that is consistent with the information reported to the offshore tax administration authorities. The FTIR data analyzed in this study consists of records that report payments of income sourced in the foreign jurisdiction when the beneficial owner of the reported income is a U.S. taxpayer.[13] The OECD describes the data as follows:

[13] Although there is no legal definition of "beneficial owner" for treaty purposes (http://www.oecd.org/tax/treaties/47643872.pdf), we treat the term as synonymous with "tax resident." To obtain treaty benefits, many U.S. treaty partners require the IRS to certify that the person claiming treaty benefits is a resident of the U.S. for Federal tax purposes. The IRS provides this residency certification on Form 6166, a letter of U.S. residency certification. The IRS procedure for requesting a certificate of residency (Form 6166) is the submission of Form 8802, *Application for United States Residency Certification*. [https://www.irs.gov/individuals/international-taxpayers/form-6166-certification-of-u-s-tax-residency]

> Information which … is typically bulk information comprising many individual cases of the same type, usually consisting of payments from and tax withheld in the supplying State, where such information is available periodically under that State's own system … {… *With the use and application of this information*} compliance is generally improved and fraud can be detected which otherwise would not have come to light. {*emphasis added*}[14]

The FTIR data has approximately 20 million records for the period 2007–2012 from the reporting countries. Each record has the payor name, the income type and amount paid, and the recipient name and type: (1) corporate; (2) individual; (3) unknown (where the recipient type is reported blank or is unknown); and (4) other (a mix of up to 20 other minor recipient categories). Payors are generally business entities operating in the source countries. We focus on payors that are affiliates of U.S. MNCs, including CFCs, DEs, and foreign corporate partnerships (FCPs).[15] The payments made to related corporate recipients are intercompany transactions. Thus, the payments should be reported as expense items on the payor's Form 5471 and as income items on the MNC's return.

Corporate-recipient records account for less than 7 percent of the 20 million records, but include almost half of the $791 trillion income reported during 2007–2012. Of the reporting countries, about two-thirds of the countries report the type of recipient, so we do not capture all corporate recipients.[16]

Table 6 reports the income types and total amounts paid to corporations versus all recipient types over 2007–2012. We focus on payments to corporations of fixed, determinable, annual, or periodic (FDAP) income—specifically interest, royalties, dividends, and capital gains.[17] The FDAP income accounts for 80 percent of the total income reported in the corporate recipient records.

TABLE 6. Corporate Recipient Income for 2007–2012

[$ amounts in millions]

Item	Corporate	All Recipients
FDAP Income		
Dividends	$175,745	$284,712
Interest	62,598	153,852
Royalties	68,489	114,287
Capital Gains	1,881	4,127
Non-FDAP Income		
Other income	46,618	158,743
Income from immovable property	966	3,175
Business profits	14,160	14,925
Missing income type	15,865	57,705
Total Income	**$386,322**	**$791,526**

NOTE: Fixed, determinable, annual, periodic (FDAP) income includes interest, royalties, dividends, and capital gains. The data includes eight other income categories that we classify as "other income."

In next section, we explain how we match recipients to the tax return data, and then identify intercompany transactions. We assume that U.S. beneficial owners of the income sourced offshore will identify themselves to the source country tax administration, usually to claim treaty preferences for tax withholding rates;[18] therefore, we should be able to match to U.S. tax records.

---

[14] OECD and Council of Europe (2011).

[15] CFCs file Form 5471; DEs appear in Form 8858 and FCPs in Form 8865.

[16] Assuming the same distribution of corporate vs. individual records (where the recipient type is known) for the records with an unknown recipient type, corporate records would constitute 70-80 percent of the total income reported.

[17] [http://www.irs.gov/Individuals/International-Taxpayers/Fixed,-Determinable,-Annual,-Periodical-(FDAP)-Income]

[18] The income recipient will claim U.S. tax residence using Form 6166 through submission of Form 8802, the application for U.S. tax residency certification. [https://www.irs.gov/individuals/international-taxpayers/form-6166-certification-of-u-s-tax-residency]

B. Matching Information Records to Tax Return Records

Comparing the FTIR information data to comparable data in the corporate e-file database requires identification of the corporate payor and recipient entities. The FTIR data does not provide a global taxpayer identification number (TIN).[19] Hence, we use quantitative linguistics to match names,[20] which involves identifying all possible entities and affiliates of consolidated corporate taxpayers that file Form 1120 tax returns with the IRS. Appendix B describes the matching process. The entities and affiliates in the e-file are defined as the participating entity list (PEL).[21] We refer to any entity name found in the e-file as a "PELname."

1. Results of Matching Process

Matched Recipients. The recipient matches to an e-file corporate entity are reported in Table 7, Panel A. The recipient matching results provide information about the amount of corporate entity income sourced offshore. Although we match only 29 percent of the recipients, these certain matches account for 69 percent of the income. Our focus in this paper is on the matched payors and their reporting of the payments as expenses on Form 5471.

Matched Payors. The payor matches are summarized in Table 7, Panel B, which shows that out of 47,076 unique payors that make payments to corporate recipients, 7,911 (17 percent) are related payors, which means that if the payor is a CFC, we should be able to match the payment to the payor's Form 5471. Certainly, payors may be affiliates of corporate taxpayers from different countries that pay income to recipients where the beneficial owner is a U.S. (corporate) taxpayer. For example, a payor could be a subsidiary of a French corporation operating in the UK, paying a U.S. entity. But some of the payors will be affiliates of U.S. MNCs (related payors), and those payors will file a U.S. tax return (Form 5471, Form 8858, or Form 8865).

These matched payors account for 55 percent of the income paid to corporate recipients. The related payors make payments to U.S. beneficial owners that average $27M ($212,000/7,911). The unrelated payor remits an average $4.5M to U.S. beneficial owners. Thus, the non-U.S. affiliated payors make lower payments to U.S. corporate recipients.[22]

Of the $212M payments from related payors, 66 percent was paid to recipients that were identified with certainty (not tabulated). We know, however, that the payments were sent to U.S. beneficial owners, whether or not we could match the recipients with certainty. Thus, the payments should be reported on the payor's tax return. If the payor is a CFC, we should be able to match the payment to the expense on the Form 5471. We use the results of the recipient process to help us identify related payors.

Table 7, Panel C. shows the entity form of the related payors. We focus on the $129B that was paid by a CFC that should be reported as an expense on a Form 5471. In particular, 80 percent of royalty payments were made by CFCs ($33,573/$42,011). DEs paid relatively more dividends (42 percent) and interest payments (46 percent); possibly reflecting the use of treaty preferences by the DE entity.

[19] If a source country TIN is reported, it may not be clear if it is a VAT number, a legal codification number (such as for incorporation), or some other designation.

[20] A virtual TIN was developed for this purpose by the IRS. The process involves the solution of a large combinatorial problem that mathematically identifies every possible combination and ordering of name-words for each payor and recipient reported in the information reporting data. See Appendix B.

[21] The PEL is a list of all business entities that are capitalized in the corporate taxpayers' commercial ventures and includes domestic and foreign corporate subsidiaries with both majority and minority positions, domestic and foreign partnerships, disregarded entities, hybrids and reverse hybrids.

[22] Recipients that may be U.S. corporations for unrelated payors may be indirectly related through a foreign parent. That is, the payor has a foreign parent, where the foreign parent is the owner of the recipient entity that is a U.S. corporation. In this instance, there is an indirect relationship between the payor and recipient that turns upon the role of the common foreign parent.

TABLE 7. Matching Results: 2007–2012

Panel A: Payments Received by Matched U.S. Corporate Recipients

Item	Certain	Uncertain	Total
N (unique)	63,921	159,311	223,232
Payments ($M)	$265,649	$117,322	$382,971
Percent of Recipients	28.6%	71.4%	
Percent of Income	69.4%	30.6%	

NOTE: The recipient matching process found multiple potential matches for 223,232 unique recipients, but there were recipients that could not be matched at all. For example, the name on the data record was simply random characters. When there was a possible match, we retained the match with the highest matching score. The certainty of the match is estimated to be 95 percent; thus 5 percent of the "certain" entity matches are not meaningful. These certain and uncertain matches account for $382,971/$386,322 (99 percent) of the payments (in $M) to U.S. corporate recipients.

Panel B: Matched Payor Payments ($M) to All U.S. Corporate Recipients

Item	Related	Unrelated	Total
N (unique payors)	7,911	39,165	47,076
Dividends	$99,910	$75,835	$175,745
Interest	36,388	26,210	62,598
Royalties	42,011	26,478	68,489
Capital Gains	182	1,699	1,881
Other	33,510	44,100	77,610
Total	**$212,001**	**$174,321**	**$386,322**
Percent of Payors	16.8%	83.2%	
Percent of Payments	54.9%	45.1%	

NOTE: Payors that are related to a U.S. MNC account for 17 percent of the payors and 55 percent of the payments to corporate recipients. The related payors make payments to U.S. beneficial owners that average $27M ($212,001/7,911) over the 2007–2012. The unrelated payors remit an average $4.5M to U.S. beneficial owners. Thus, the non-U.S. affiliated payors pay less to U.S. corporate recipients.

Panel C: Related Payor Payments ($M), by Type of Payor Entity

Item	Total	CFC	DE	FCP
Dividends	$99,910	$55,986	$42,383	$1,541
Interest	36,388	17,837	16,767	1,783
Royalties	42,011	33,573	7,628	809
Capital Gains	182	182		
Other	33,510	21,796	9,593	2,121
Total	**$212,001**	**$129,374**	**$76,372**	**$6,255**

2. CFC e-File Details

This section discusses the e-file part of the matching process and uses royalty expense to illustrate. Although we focus on royalty payments, we provide other detail from Form 5471 as background information. Table 8, Panel A, shows the total amounts of the following Form 5471 line items for all CFCs in the reporting and non-reporting countries for 2007–2012:

- CFC income (per books) from Schedule C, line 21;
- Assets (end of period) from Schedule F line 22b;
- Royalty income from Schedule C, line 6b;
- Royalty expense from Schedule C, 11b; and
- Intangible assets (end of period) from Schedule F line 11a+11b+11c.

Relative percentages of global totals for each of these line items is presented in Panel B. For example, in 2007, the reporting countries account for 51 percent of total CFC income, 59 percent of assets, 24 percent of the royalty income, 42 percent of the royalty expense, and 60 percent of intangible assets.

TABLE 8. CFC Form 5471 by Reporting/Nonreporting Countries: 2007–2012
[$ amounts in millions]

Panel A: Form 5471 Line Items by CFCs in Reporting vs. Nonreporting Countries

Category	Item	2007	2008	2009	2010	2011	2012
Reporting	**N**	**41,387**	**43,298**	**43,137**	**43,441**	**45,150**	**45,969**
	CFC income	1,048,722	1,153,273	958,128	1,066,877	1,349,711	1,139,497
	Total assets	8,774,534	8,775,613	6,829,747	7,861,438	8,374,122	8,744,304
	Royalty income	4,753	9,802	9,455	10,780	13,357	11,506
	Royalty expense	15,985	23,245	20,603	20,900	34,417	24,591
	Intangibles	293,831	281,047	325,136	341,586	413,109	405,353
Nonreporting	**N**	**25,055**	**26,099**	**26,559**	**27,557**	**29,144**	**30,063**
	CFC income	988,070	1,104,552	1,078,703	1,346,182	1,304,061	1,541,212
	Total assets	6,136,057	6,400,989	6,840,861	7,558,854	8,282,967	9,531,800
	Royalty income	14,992	26,338	23,841	21,974	37,952	36,974
	Royalty expense	18,713	31,613	30,616	36,006	39,824	47,101
	Intangibles	194,833	202,081	227,007	256,374	310,830	376,609
Other	**N**	**401**	**303**	**162**	**147**	**182**	**194**
	CFC income	21,516	19,985	12,865	14,335	22,254	18,188
	Total assets	56,181	44,739	29,917	33,816	49,730	71,374
	Royalty income	52	1	0	1	53	34
	Royalty expense	3,561	40	13	11	4	7
	Intangibles	1,264	829	322	523	550	1,246
Total	**N**	**66,843**	**69,700**	**69,858**	**71,145**	**74,476**	**76,226**
	CFC income	2,058,308	2,277,811	2,049,696	2,427,394	2,676,025	2,698,897
	Total assets	14,966,771	15,221,341	13,700,525	15,454,107	16,706,819	18,347,478
	Royalty income	19,797	36,141	33,297	32,754	51,362	48,515
	Royalty expense	38,259	54,898	51,233	56,917	74,245	71,699
	Intangibles	489,928	483,957	552,466	598,484	724,489	783,209

Panel B: Reporting Countries' Share of Total CFC Line Items

Category	Item	2007	2008	2009	2010	2011	2012
Reporting	N	62%	62%	62%	61%	61%	60%
	CFC income	51%	51%	47%	44%	50%	42%
	Total assets	59%	58%	50%	51%	50%	48%
	Royalty income	24%	27%	28%	33%	26%	24%
	Royalty expense	42%	42%	40%	37%	46%	34%
	Intangibles	60%	58%	59%	57%	57%	52%

IV. Tax Compliance Inferences From Matched FTIR Data

A. FTIRC Measures

We create two FTIRC measures that are consistent with conventions used in the estimation of the tax gap. The IRS tax gap estimation methodology defines two components of the tax gap:

- **nonfiling gap**—the tax not paid on time by those who do not file tax returns on time; and
- **underreporting gap**—the tax misreported on timely filed returns.

The recipient filing measure (FTIRC1) provides a means to quantify the nonfiling gap, albeit for that portion of the filers that identify themselves as U.S. beneficial owners to a foreign tax administration (FTA). We assume that if the corporate recipient entity reported in the FTIR data takes on the burden of doing so, then that entity should also be reported in the corporate e-file.

The expense reporting ratio (FTIRC2) is characterized as quantifying elements of income underreporting. That is, FTIRC2 is the ratio of payments reported to the FTA relative to the expense reported to the IRS. IRS research has shown that accurate third-party information reporting is a significant determinant of individual income tax reporting compliance. Thus, when we draw compliance inferences using FTIRC2, we assume that accurate third-party reporting of offshore corporate tax positions should also be consistent with corporate tax compliance, especially for offshore income-shifting.

Drawing inferences about compliant tax positions is a conceptual departure from most antecedent tax compliance research that uses evidence from noncompliant taxpayers to draw inferences about tax compliance in the population. We draw on the CAP/non-CAP income-shifting comparisons (see section II-C3) that suggests CAP taxpayers have more tax compliant positions vis-à-vis the positions of corporate taxpayers that are included in the non-CAP sample; but more importantly, CAP firms are very different (i.e., have tax positions that are more compliant) from the population of corporate taxpayers in general.

The two FTIR compliance measures reflect how accurately CFC recipient entities: (1) file returns with the IRS (when the recipient entity is a U.S. tax resident); and (2) report FDAP payments as an expense on Form 5471, Schedule M, Related Transactions Statement, relative to what the CFC reports to the FTA. Thus, we compare and contrast what the payor reports to the foreign tax administration with what the payor reports to the IRS. Then, using the FTIRC2 measure that quantifies the comparisons of the FTA tax data to the IRS tax data, inferences are drawn about compliant income-shifting.

1. FTIRC1 Measure—Recipient Filing

Measurement: We first examine whether the recipient entities of FDAP file a U.S. tax return. The foreign payor provides the FTA information similar to what U.S. payors of FDAP provide to the IRS using Form 1042-S: the recipient, amount, and type of payments. We rely on the virtual TIN from the matching process described in Part III in which recipients are either matched with certainty or do not have a meaningful match to an entity in the corporate e-file.[23] The match logic supporting the virtual TIN assumes that if the recipient has identified itself as a U.S. tax resident (an entity with a beneficial owner that is a U.S. corporate taxpayer) to the FTA, the recipient will also identify itself to the IRS in the corporate e-file. Thus, unmatched recipients suggest potential filing noncompliance: the recipient is either not identifying itself to the IRS or is not identifying itself correctly to the FTA. For example, the recipient may notify the FTA that it is a U.S. tax resident (or there is a U.S. corporate beneficial owner) to claim beneficial tax treaty withholding tax rates, or the payor is providing poor recipient identification to the tax authorities offshore.

FTIRC1 is the ratio of the matched recipients to all recipients for each related payor, weighted by the FDAP payments. Thus, FTIRC1 is the percent of the payments reported to the FTA for recipients matched to the corporate e-file. For example, suppose a related payor has 10 recipients in the FTIR data, and each recipient receives a payment of $100. Six of the recipients are matched to the corporate e-file, and four are unmatched. The FTIRC1 ratio is 60 percent (600/1,000).

As shown in Table 9, the average CAP FTIRC1 measure is 88 percent, meaning 88 percent of all FDAP payments were made to recipients identified in the e-file. The non-CAP average is 75 percent. The difference between the FTIRC1 measure for CAPs and non-CAPs is significant at the 95 percent level of certainty. Thus, CAP-related payors report more matched recipients relative to non-CAP-related payors, suggesting that CAP recipient entities have better filing compliance with the IRS.

[23] Type II errors are also possible if the matching process failed to identify a recipient that does exist in the e-file.

TABLE 9. FTIRC1

Item	CAP	Non-CAP	All Others	Total
N	224	418	3,460	4,102
FTRIC1	0.883	0.754	0.670	0.690

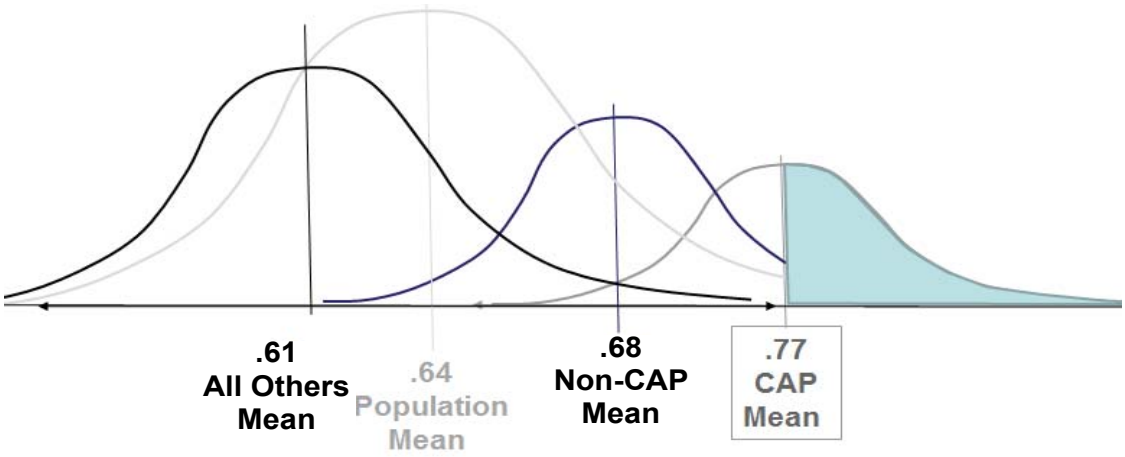
NOTES: FTIRC1 is the ratio of the matched recipients to all recipients for each related payor, weighted by the FDAP payments. Thus, FTIRC1 is the percent of the payments reported to the FTA for recipients matched to the corporate e-file. As shown in Table 9, the average CAP FTIRC1 measure is 88 percent, meaning 88 percent of all FDAP payments were made to recipients identified in the e-file. The non-CAP average is 75 percent. The difference between the FTIRC1 measure for CAPs and non-CAPs is significant at the 95 percent level of certainty. Thus, CAP-related payors report more matched recipients relative to non-CAP-related payors, suggesting that CAP recipient entities have better filing compliance with the IRS.

This table is based on an earlier FTIR dataset for the years 2007–2010. The "All Others" are Large Business & International (LB&I) taxpayers that were not in the CAP/non-CAP sample. We will update the table using the dataset used for the other tables.

We infer that when a CFC's suppliers are U.S. beneficial owners filing with the IRS, the CFC is engaged in intercompany financial transactions that are reported to the IRS. Recipients that claim to be U.S tax residents (are identified as U.S. tax residents) to the FTA in the source country but are not identified in the e-file merit further review.

The compliance inference that is drawn using the FTIRC1 measure is illustrated in Figure 1. The sample means of FTIRC1 for CAP MNCs, matched non-CAP MNCs, and all other corporate taxpayers are significantly different, meaning it is unlikely these separate samples could be drawn from the population at random. Thus, the FTIRC1 measure reflects fundamental differences among these three groups.[24]

FIGURE 1. FTIRC1—Recipient Filing



CAPs, non-CAPs, and All Others have FTIRC1 means that are significantly different. Thus, the groups are not likely to be drawn from the same population. The inference for CAPs is that 77 percent of cross-border FDAP payments are with entities known to the IRS.

2. FTIRC2 Measure—Expense Reporting

The second measure is based on the assumption that accurate third-party reporting is coincident with corporate tax compliance—a relationship demonstrated for individual taxpayer compliance. FTIRC2 is computed for each matched payor for each year as the ratio of expense reported by the CFC to the FTA, relative to the expense reported by the CFC to the IRS. The FTA numerator uses all recipient royalties and interest reported in the FTIR data, whether or not paid to a related party or even matched to an e-file recipient.[25] The IRS

[24] If the population means are not different, then the inference process must be applied for each entity as we do when we compute FTIRC2. That is, each CFC's FTIRC value is compared to the mean for that group, using a 1-tail test.

[25] An alternative measure for FTIRC2 can be computed using data for the FDAP reported to the FTA that is ascribed with certainty to entities in the e-file. This FTIRC2 measure would be lower.

denominator includes royalty and interest expense reported to the IRS on the payor's Form 5471 [Schedule C, *Income Statement* for royalty expense (line 11b) and interest expense (line 12)].[26] If the CFC paid all of its royalties and interest to U.S. beneficial owners, the ratio of expense reported to the FTA should equal the expense reported to the IRS, and FTIRC2 would be equal to 1. CFCs, however, also make payments to entities that are not U.S. beneficial owners, in which case the FTIRC2 will be less than 1.[27]

We evaluate FTIRC2 on a group basis, comparing the payor's ratio to the group's mean ratio computed across years. The group is based on industry and excludes CFCs that own a DE. The FTIRC2 measure could be affected if the CFC makes payments to a related DE. Such payments would be reported to the FTA, but not reported on the Form 5471 because DE payments are eliminated as intracompany transactions.[28]

The average FTIRC2 for royalty expense reported by CAP MNCs and non-CAP MNCs is shown in Table 10, Panel A. For example, the average FTIRC2 for CAPs is 25 percent in 2010. Thus, on average, one quarter of the CFCs' royalty remittances reported to the IRS are also reported to the FTA in the FTIR data. The corresponding FTIRC2 for non-CAP MNCs is 33 percent in 2010.

The average FTIRC2 for interest expense is shown in Table 10, Panel B. For example, the FTIRC2 for CAP MNCs is 1.4 percent in 2007. The corresponding FTIRC2 for non-CAP MNCs is 2.9 percent. Evidently, most CFC interest payments are not paid to U.S. tax residents because only a small percentage of all interest expense reported by CFCs to the IRS are also reported to the FTA in the FTIR data. Additional work is ongoing to develop the compliance measures further. Any results reported in this paper are preliminary.

TABLE 10. FTIRC2—Royalty and Interest Expense, 2007–2012

Panel A: Mean FTIRC2—Royalty Expense

Year	CAP	Non-CAP	All Others
2007	0.192	0.396	0.240
2008	0.240	0.444	0.280
2009	0.262	0.319	0.241
2010	0.248	0.326	0.290
2011	0.247	0.317	0.291
2012		0.402	0.271

Panel B: Mean FTIRC2—Interest Expense

Year	CAP	Non-CAP	All Others
2007	0.014	0.029	0.046
2008	0.011	0.033	0.040
2009	0.032	0.023	0.044
2010	0.020	0.017	0.034
2011	0.023	0.056	0.040
2012		0.009	0.036

NOTE: FTIRC2 is computed for each matched payor for each year as the ratio of payments reported by the CFC to the FTA, relative to the expense reported by the CFC to the IRS on Form 5471 [Schedule C, *Income Statement* for royalty expense (line 11b) and interest expense (line 12). If the CFC paid all its royalties and interest to U.S. beneficial owners, the ratio of expense reported to the FTA should equal the expense reported to the IRS and FTIRC2 would be 1. CFCs, however, also expense payments to entities that are not U.S. beneficial owners, in which case the FTIRC2 will be less than 1.

[26] In future work, we plan to also use Schedule M, *Transactions Between Related Entities*, for royalties paid (line 20, cols. b–f) and interest paid (line 22, cols. b-f).

[27] We determine the measure only for CFC payors filing Form 5471. We could also compute the measure for DE and FCP payors. However, neither the DE, Form 8858, nor the FCP, Form 8865, include a royalty expense line item detail in the reported income statement amounts. Future work could use expenses reported on Form 8858, Schedule M, *Transactions Between Foreign Disregarded Entity of a Foreign Tax Owner and the Filer or Other Related Entities*, and Form 8865, Schedule N, *Transactions Between Controlled Foreign Partnership and Partners or Other Related Entities.*

[28] If the FTIRC2 observations for CFCs with DEs were included with CFCs that do not have DEs, the income consolidation rules imply that the reported FTA expense>IRS expense. Thus, for CFCs with DEs, the permissible range of FTIRC2>1, whereas the range for CFCs without DEs is between 0 and 1. If these two distributions are combined, compliance inferences would be confounded because the two ranges are discrete.

B. Compliant IS

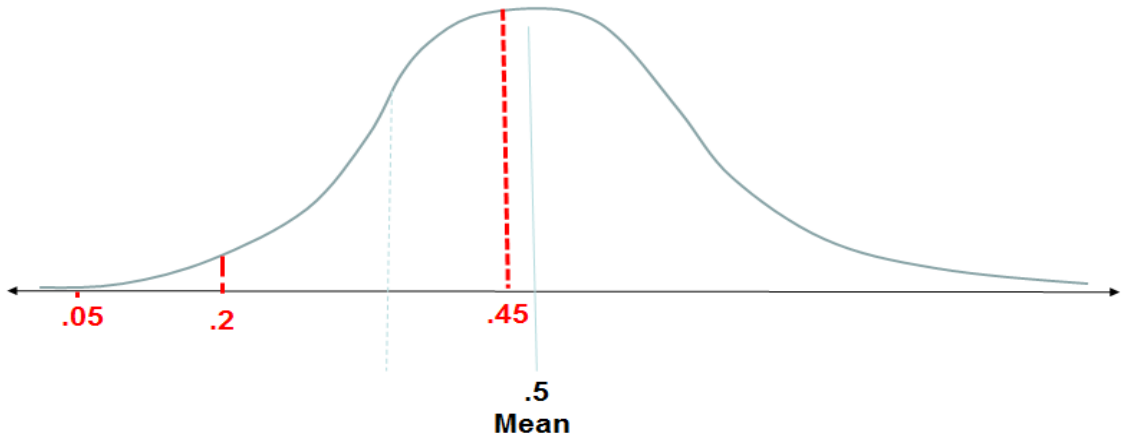
This section describes the process we use to draw compliance inferences about the estimated levels of income-shifting (IS) by CFCs. The ISM provides income-shifting estimates for each CFC, but does not identify whether the income-shifting estimate is compliant or noncompliant. The inferences that are drawn about the compliant portion of the estimated income-shifting levels rely upon the cumulative probability distribution of the CFCs' FTIRC2 measures. While the income-shifting estimates are determined for all CFCs in both reporting and nonreporting countries, the inferences about compliant income-shifting may be determined only for CFCs domiciled in reporting countries. We assume that the FDAP "footprint" depicted by the FTIR data provides an empirical proxy measure for the compliance of the estimated income-shifting.

We reject an inference of compliance when FTIRC2 (FTA expense/IRS expense) of the given CFC is significantly different from the ratio observed in the group of like CFCs (based on industry and DE). We use the FTIRC2 measures to compute a compliance adjustment factor (CAF) at the CFC level. If a CFC's FTIRC2 < mean FTIRC2 for the CFC's group, we reduce the estimated income-shifting amount by (1 – CAF) where CAF = FTIRC2/mean FTIRC2. CAF is first computed for royalties using the royalty FTIRC2 and then for interest expense using the interest FTIRC2. We then compute a weighted average of the two CAFs, where the weights are the relative share of (interest or royalty) expense in the total of interest expense and royalty expense in Schedule C. For example, suppose for a CFC, that Schedule C reports a royalty expense of $60 and an interest expense of $40, and that the royalty CAF measure is 0.4, with the interest CAF measure equal to 0.3. If so, then the overall CAF applied to the income-shifting estimate is 0.36 (= 0.6*0.4+0.4*0.3).

The CAF derivation relies upon the cumulative distribution of the FTIRC2 measure. For example, assume that a CFC has an IS of $100. Further assume the CFC's royalty FTIRC2 is 0.2 and for simplicity no interest expense is reported by the CFC to the IRS. Thus, what the CFC reported to the FTA is only 20 percent of what is reported to the IRS. Suppose this CFC is in a group of comparable CFCs with a mean FTIRC2 of 0.5. Thus, because the FTIEC2 is below the mean, the CAF is 0.2/0.5, or 40 percent.

The example is illustrated in Figure 2. When the FTIRC2 measure is less than the group's mean, the CFC has a ratio of FTA/IRS expense that is less than its peers. Continuing the previous example, the CFC's group reports royalty expense to the FTA that is on average 50 percent of what the group reports to the IRS. The CFC with royalty FTIRC2=0.2, however, reports to the FTA only 20 percent of the expense it reports to the IRS. Thus, this CFC is conducting transactions that are at variance from the group CFCs with the result that the compliance adjustment factor for this CFC based on FTIRC2 is 40 percent, which suggests that the CFCs reported financial transactions are only 40 percent of what is expected for like CFCs.

FIGURE 2. FTIRC2—Expense Reporting



NOTES:

FTIRC2 = .5. The group's average royalty expense reported to the IRS is 50 percent of what is reported in the source country to the FTA. ISp is deemed 100 percent compliant when FTIRC2 >=.5.

FTIRC2 = .45. The CFC's royalty expense reported to the IRS is 45 percent of what is reported in the source country. Because this CFC's FTIRC2 is below the mean, 90 percent (.45/.5) of the ISp is deemed to be compliant.

FTIRC2 = .2. 40 percent (.2/.5) of the ISp is deemed to be compliant.

FTIRC2 = .05. 10 percent (.05/.5) of the ISp is deemed to be compliant.

Table 11 presents the levels of estimated income-shifting that reflect compliant offshore positions of the CFC, where the CFC is in an information reporting country. When a CFC does not have an FTIR record, the mean value of the FTRIC2 measure is assigned to that CFC. The table reports the results for CAP MNCs, non-CAP MNCs, and all other CFCs. The differences are not significant for CAPs, non-CAP, and others. In general, we infer that just under half of the ISp is "compliant." We caution that these results are preliminary, and our research in ongoing.

TABLE 11. Compliant Income-Shifting Estimates in FTIR Countries Based on FTIRC2 for IS>0

Year	IS			Compliant IS			Compliance Adjustment Factor (Percent Compliant)		
	CAP	Non-CAP	Others	CAP	Non-CAP	Others	CAP	Non-CAP	Others
2007	11,085	14,852	48,029	4,842	6,611	21,106	0.437	0.445	0.439
2008	6,287	12,085	60,905	2,962	5,471	26,723	0.471	0.453	0.439
2009	11,275	12,874	55,309	5,338	5,894	24,789	0.473	0.458	0.448
2010	10,113	15,814	61,709	4,785	7,107	28,029	0.473	0.449	0.454
2011	11,613	18,572	68,353	5,326	8,287	29,870	0.459	0.446	0.437
2012	10,590	18,273	48,595	4,973	8,223	21,487	0.470	0.450	0.442

NOTE: We evaluate FTIRC2 on a group basis, comparing the payor's ratio to the group's mean ratio computed across years. The group is based on industry and DE. When a CFC does not have an FTIR record, the mean value of the FTRIC2 measure for that group is assigned to that CFC. We use the FTIRC2 measures to compute a compliance adjustment factor (CAF) at the CFC level. If a CFC's FTIRC2 < mean FTIRC2 for the CFC's group, we reduce the estimated income-shifting amount by (1-CAF) where CAF = FTIRC2/mean FTIRC2. CAF is first computed for royalties using the royalty FTIRC2 and then for interest expense using the interest FTIRC2. We then compute a weighted average of the two CAFs, where the weights are the relative share of (interest or royalty) expense in the total of interest expense and royalty expense in Schedule C. For example, suppose that for a CFC, Schedule C reports royalty expense is $60 and interest expense is $40, and that the royalty CAF measure is .4, with the interest CAF measure equal to .3. The overall CAF applied to the income-shifting estimate is .36 (60*.4+.40*.3).

V. Conclusion

This exploratory research illustrates a potential use of FTIR data to identify compliant income-shifting. We first extend the prior income-shifting model enabling estimation of income-shifting on a CFC/country-by-country basis (N-state model). We then match the FTIR data to e-file data, which allows us to use third-party reporting to examine compliance. We construct two FTIR compliance measures that have foundations in tax compliance related to the filing gap and the income underreporting gap. Whether the FTIRC measures could be useful for tax gap estimation remains open to further study by using enforcement data. The initial specifications are designed to analyze taxpayer compliance and make use of the distinctions between CAP and non-CAP taxpayers to demonstrate the inference framework.

Evidence provided by the FTIRC measures confirms some but not all expectations about taxpayer compliance. First, the income-shifting results suggest that CAP taxpayers have less estimated income-shifting. Second, the recipient entity filing measure (FTIRC1) showed expected differences between CAP, non-CAP and all other filers. The last measure (FTIRC2), however, did not support the conclusion that CAP taxpayers are more compliant. Further study is needed to confirm the attributes used to create the groups.

Our data has measurement errors due to data quality problems in both the FTIR data and the corporate e-file data. Observations from the e-file that have data quality problems have been dropped. Had they not, the ISM solution would require notional adjustments in line items reported by the taxpayer in either the numerator or denominator to reach a reasonable ATROC. The most serious e-file data exceptions have dramatic impacts on the shifting estimates from the ISM. These dropped entities contain CFCs with asset valuation issues, especially for what we expect are intangible assets and may need further scrutiny to address the valuation reporting weaknesses.

We recognize the effects of the asset-valuation problem as being one of the determinants of the data exceptions in the shifting estimates. As noted, CFCs with about half of the offshore income (where CFC income is positive, i.e. excluding CFC NOLs) are affected by the asset-valuation issues. We surmise these asset-valuation issues are likely to have emerged from intangible assets being expensed during development and within cost sharing platforms. Since the income-shifting model depends upon accurate asset valuations, this measurement problem implies that there would be significant limitations for evaluating arms' length transfer pricing since

asset values are pivotal data inputs for transfer pricing computations. Hence, we conclude that intangible expensing creates statistical irregularities for tax administration that severely limit effective enforcement of the arms' length standard using the income-shifting concepts analyzed here, as well as for transfer pricing methods in general. This finding is consistent with conclusions of a National Academy of Sciences Workshop[29] that concluded intangible expensing created challenges for the accurate measurement of investment in financial statements and in the national income and product accounts (NIPA). These investment measures influence Wall Street financial decisions, and are a key data input used by the Federal Reserve.

Perhaps the largest hurdle we had to overcome in this research is creating a virtual TIN using computational linguistics to match FTIR data to e-file data. In the process of recipients confirming to the FTA that they are U.S. residents, the U.S. TIN (EIN for corporations) should be included in the shared FTIR data. The new country-by-country reporting initiative may resolve many, but not all, of the entity identification issues.

[29] National Research Council (2009).

References

Altshuler, R., and H. Grubert. 2005. The Three Parties in the Race to the Bottom: Host Governments, Home Governments, and Multinational Companies. *Florida Tax Journal* 7(3): 154–183.

Black, T. 2012. Quantification of Taxpayer Text Identifiers: ALPHA Function, IRS Office of Research, February, Washington, DC.

Christian, C., and T. Schultz. 2005. ROA-Based Estimates of Income-Shifting by U.S. Multinational Corporations. IRS Research Conference, 2005, Washington, DC.

Collins, J., D. Kemsley, and M. Lang. 1998. Cross-Jurisdictional Income-Shifting and Earnings Valuation. *Journal of Accounting Research* 36(2): 209–229.

Desai, M. A., C. F. Foley, and J. R. Hines Jr. 2006. The Demand for Tax Haven Operations. *Journal of Public Economics* 90(3): 513–531.

Dharmapala, D., and J. R. Hines Jr. 2009. Which Countries Become Tax Havens? *Journal of Public Economics* 93(9–10): 1058–1068.

Dharmapala, D., and N. Riedel. 2013. Earnings Shocks and Tax-Motivated Income-Shifting: Evidence from European Multinationals. *Journal of Public Economics* 97: 95–107.

Dowd, T., P. Landefeld, and A. Moore. 2017. Profit Shifting of U.S. Multinationals. *Journal of Public Economics* 148: 95–107.

Dyreng, S.D., and B.P. Lindsey. 2009. Using Financial Accounting Data To Examine the Effect of Foreign Operations Located in Tax Havens and Other Countries on U.S. Multinational Firms' Tax Rates. *Journal of Accounting Research* 47(5): 1283–1316.

Dyreng, S. D., and K. S. Markle. 2016. The Effect of Financial Constraints on Income-Shifting by U.S. Multinationals. *Accounting Review* 91(6): 1601–1627.

Grubert, H. G. 2012. Foreign Taxes and the Growing Share of U.S. Multinational Company Income Abroad: Profits, not sales, are being globalized. *National Tax Journal* 65(2): 247–275.

Grubert, H. G. 1998. Taxes and the Division of Foreign Operating Income Among Royalties, Interest, Dividends and Retained Earnings. *Journal of Public Economics* 68(2): 269–290.

Hines, J. R., and E. M. Rice. 1994. Fiscal Paradise: Foreign Tax Havens and American Business. *The Quarterly Journal of Economics* 109(1): 149–182.

Jenkins, G. P., and B. D. Wright. 1975. Taxation of Income of Multinational Corporations: The Case of the United States Petroleum Industry. *Review of Economics and Statistics*. 57(1):1–11.

Kopits, G. F. 1976. Taxation and Multinational Firm Behavior: A Critical Survey, *IMF Staff Papers* 23 (3):513–516.

Klassen, K., and S. Laplante. 2012. Are U.S. Multinational Corporations Becoming More Aggressive Income Shifters? *Journal of Accounting Research* 50 (5): 1245–1285.

Manning, C. D., P. Raghavan, and H. Schütze. 2008. *Introduction to Information Retrieval*, Cambridge University Press. [http://www-nlp.stanford.edu/IR-book/].

Markle, K. 2015. A Comparison of the Tax-Motivated Income-Shifting of Multinationals in Territorial and Worldwide Countries. *Contemporary Accounting Research* 22(1): 7–43.

National Research Council. 2009. Intangible Assets: Measuring and Enhancing Their Contribution to Corporate Value and Economic Growth. Washington, DC: The National Academies Press. https://doi.org/10.17226/12745.

OECD. 2013. Action Plan on Base Erosion and Profit Shifting, OECD Publishing. http://dx.doi.org/10.1787/9789264202719-en.

OECD and Council of Europe. 2011. The Multilateral Convention on Mutual Administrative Assistance in Tax Matters: Amended by the 2010 Protocol, OECD Publishing. http://dx/doi.org/10.1787/9789264115606-en.

Rousslang, D. 1997. International Income-Shifting by U.S. Multinational Corporations. *Applied Economics* 29 (71): 925–934.

Salton, G., A. Wong, and C. S. Yang. 1975. A Vector Space Model for Automatic Indexing, *Communications of the ACM*, vol. 18, nr. 11, pages 613–620. [http://www.cs.uiuc.edu/class/fa05/cs511/Spring05/other_papers/p613-salton.pdf].

Wilde, J., and R. Wilson. 2017. Perspectives on Corporate Tax Avoidance: Observations from the Past Decade. Working Paper.

Appendix A

Derivation of the N-State Income-Shifting Model

The income-shifting framework provides a methodology for quantifying the income-shifting activities of U.S. multinational corporations based on the theoretical work of Rousslang (1997). The framework uses a profit optimization model that assumes firms account for the benefits of income-shifting when making investment decisions and, consistent with economic theory, allocate capital so that marginal after-tax rates of return on assets are the same across all members of an affiliated group. Factor price equalization in a single factor context implies that capital is allocated among jurisdictions to realize the greatest overall return. Capital is reallocated from jurisdictions where the returns are lower and moved into jurisdictions where returns are higher, resulting in factor price equalization.[30]

Specifically, the original Christian and Schultz[31] income-shifting two-state solution yields an annual estimate of the direction and magnitude of income shifted between a firm's domestic and foreign operations as a function of the pretax income, tax, and long-term asset values reported in tax records. Because the two-state (domestic and foreign) solution aggregates the rate of return in all observed foreign jurisdictions, it does not distinguish between the amounts of income-shifting attributed to individual foreign tax jurisdictions. The IRS extended the two-state income-shifting framework to reflect a three-state solution (domestic, foreign #1, and foreign #2), so the amounts of income shifted between a firm's domestic operations and each of two foreign super-jurisdictions can be estimated.

The income-shifting framework posits convergence to a single capital factor price across jurisdictions.

If there is equivalence between all three jurisdictions in the rate of return after tax when adjusting for shifting, then:

$$\frac{[Y_d + IS_T] * (1-t_d)}{K_d} = \frac{[Y_1 - IS_1] * (1-t_1)}{K_1} = \frac{[Y_2 - IS_2] * (1-t_2)}{K_2},$$

where:

Y_d = pretax domestic income

Y_1 = pretax foreign jurisdiction 1 income

Y_2 = pretax foreign jurisdiction 2 income

K_d = domestic capital stock

K_1 = foreign jurisdiction 1 capital stock

K_2 = foreign jurisdiction 2 capital stock

t_d = tax rate domestic (T_d/Y_d) where T = level of taxes

t_1 = tax rate foreign jurisdiction 1

t_2 = tax rate foreign jurisdiction 2

IS_T = total income shift [$IS_T = IS_1 + IS_2$]

IS_1 = income shift jurisdiction 1

IS_2 = income shift jurisdiction 2.

[30] While the ISM does not provide compliance inferences for the tax transactions to shift income, it does provide estimates of the aggregate level of profits shifting offshore. The actual rate of return realized by some firms will be higher or lower than what the model predicts in a given jurisdiction, but on average and in the aggregate, the model's estimate of the total level of income shifted is likely to approximate what is actually realized. Firms expect higher-than-normal returns in foreign jurisdictions to compensate for the added risks. The ISM accounts for the risk in an underwriting framework and aggregates the total returns across all firms to provide accurate income shifting estimates.

[31] Christian, Charles and Thomas Schultz, "ROA-Based Estimates of Income Shifting by U.S. Multinational Corporations," IRS Research Conference, 2005, Washington, DC.

This paper extends the three-state model to an N-state model as derived below. Assuming factor-price equalization in a single factor framework, it follows that the capital price will converge across jurisdictions for each corporation.[32] That is, a corporation will allocate capital resources such that the rate of return to capital will be equal between any two jurisdictions (jurisdiction "i" and jurisdiction "j"). In general form, showing the effect of income shifted between the jurisdictions is as follows:

$$(Y_i - X_i) * \alpha_i = (Y_j - X_j) * \alpha_j \quad [1A]$$

where:

$$\alpha_i = (1 - t_i) / K_i \quad [1B]$$

$$\alpha_j = (1 - t_j) / K_j. \quad [1C]$$

Then, solving for the level of corporate income shifted in jurisdiction "i",

$$X_i = Y_i - (Y_j - X_j) * \alpha_j / \alpha_i. \quad [1D]$$

Now, including the domestic jurisdiction in the expression, the rate of return in the domestic jurisdiction is equal to the rate of return in any jurisdiction "j". That is, when accounting for the corporate income shifted offshore, and adjusting both the domestic and offshore jurisdictions for the corporate income shifted out of the domestic jurisdiction and into the offshore jurisdiction, we obtain:

$$(Y_d - X_d) * \alpha_d = (Y_j - X_j) * \alpha_j \quad [1E]$$

where:

$$\alpha_d = (1 - t_d) / K_d, \quad [1F]$$

and where the total income shifted out of the domestic jurisdiction (IS or Xd) is equal to the sum of the (negative) income shifted in each of the "N" offshore jurisdictions (and the domestic jurisdiction "d" is defined for j=1):

$$IS = X_d = - \sum_{j=2}^{N} X_j. \quad [1G]$$

The domestic pre-tax corporate income plus the corporate income shifted is equal to the terms on the right hand side of equation [2] below:

$$Y_d + \sum_{j=2}^{N} X_j = (Y_j - X_j) * \alpha_j / \alpha_d. \quad [2]$$

Solving equation [2] for the income-shifting term (Xj) provides the general form of the income-shifting in any jurisdiction "j" where X_1 is X_d (the total shifting from/to the domestic jurisdiction) and $X_1 = X_d = - \sum_{i=2}^{N} X_i$. Then,

$$X_2 = [(\alpha_2 + \Sigma_{i=3}^{N} \alpha_1,\alpha_2/\alpha_i) * Y_2 - \alpha_1 * Y_1 - \alpha_1 * \Sigma_{i=3}^{N} Y_i] / (\alpha_1 + \alpha_2 + \Sigma_{i=3}^{N} \alpha_1,\alpha_2/\alpha_i) \quad [3A]$$

and for any CFC in the jurisdiction "j":

$$X_j = Y_j - \alpha_2/\alpha_j * Y_2 + \alpha_2/\alpha_j * X_2, \quad [3B]$$

[32] We assume that the shifting of income from/to jurisdiction $\mathbf{X_i}$ and $\mathbf{X_j}$ will be reflexive; meaning that $\mathbf{X_i} = - \mathbf{X_j}$ such that the income shifted out of one jurisdiction must equal the income shifted into the other jurisdiction in the 2-state case. The N-state condition is found in expression 1G.

where the total income-shifting to/from the domestic jurisdiction (X_d=X_1) is:

$$X_1 = X_d = -\sum_{i=2}^{N} X_i \quad . \qquad [3C]$$

Equation [3C] shows the general form of the solution for the level of corporate income shifted into (or out of) any jurisdiction "j". The sign of the data in the solution for Xj reflects whether the income enters or leaves the jurisdiction. If the sign is negative, income is shifted into the jurisdiction. If the sign is positive, then income is shifted from this jurisdiction.

Aggregating the income shifted across all jurisdictions, the total level of income-shifting (IS) in the domestic jurisdiction (restating equation 3C) is:

$$\text{IS: } X_d = -\sum_{i=2}^{N} X_i \quad . \qquad [4]$$

Appendix B

Matching Foreign Tax Information Reporting Data to the Corporate e-File Using Computational Linguistics

This appendix provides an overview of a process to match foreign tax information reporting (FTIR) records to IRS corporate tax return records. There is no global TIN available, and thus the matching of payor and recipients to a corporate tax return is done using name matching.

The corporate tax return names are drawn from the participating entity list (PEL). The PEL is a list of all business entities in the corporate taxpayers' commercial ventures and includes domestic and foreign corporate subsidiaries with both majority and minority positions, domestic and foreign partnerships, disregarded entities, hybrids and reverse hybrids. The tax return names and the information reporting names are subject to pre-process to recognize the data quality in the entity name and provide greater uniformity in the text presentation and text format. Two different areas of data quality affect the matching process: blank names and semblants. Blank recipient names and blank payor names are found in the information reporting records. When an information reporting record has a blank payor or recipient name, there is no match logic that can be applied, and the record is deleted.

Semblants are minor differences in entity names that reflect the inconsistent use of capitalization, punctuation, and word abbreviations. Semblants can also reflect misspellings of words, but only minor differences in capitalization, punctuation and abbreviations are included.

A. Recipient Matching Methods

Recipient names are matched to PEL names using a two-step process. The first step finds all possible match candidates and the second step assigns a rank or certainty to select the "best" matching candidate name. All possible candidate names are determined by examining every combination of the name-words in a name reported in the information reporting data and comparing them to every combination of name-words in the list of PELnames. The ranking or certainty of the "best" match, among all the PELname candidates for a name in the information reporting data, is computed using the vector scoring model.[33]

The number of combinations is vast and a direct solution, representing a match, is computed using an analytical approach from quantitative linguistics that develops numeralizations of name-words using integer mapping of character sets. Each word in the name from the information reporting data and PELname is converted using a numeralization into a unique (and invertible) number. These name-word numeralizations are input to a mathematical function to compute a very large integer.[34] This very large integer number is essentially unique to each name.[35] A PELname candidate is found for a recipient name using the numerical solution where the functional value for the word combinations of a PELname candidate is equal to the functional value for the word combinations of a name reported in the information reporting data. The existence of a solution implies that one or more name-words in the PELname candidate are common to the name-word(s) in the name reported in the FTIR data.

The PELname candidate name-word combinations for the name-word combinations of a name reported in the FTIR record identifies name-words that match for some of the words in the names but, not necessarily all the name-words. That is, some but not all of a PELname candidates' name-word may match only some of the name-words in a name in the information reporting data record. In other words, it is a partial match. The numeralizations of the name-words ensure that any ordering of a given combination of name-words can be found in the functional solution. For example, suppose a name is composed of three words (A B C with sepa-

[33] See Manning, *et al.* (2008).

[34] The functional value is the multiplicative product of the numbers corresponding to the name-words. Each name-word number represents the numeralizations of each name-word. The functional value is a very large number and numerical precision is needed for all digits to make comparisons for matching. The functional values have about 100 to 150 digits (or more), and keeping computational precision requires special software routines to increase the number of registers used in the computations as most software programming languages lose mathematical precision for numbers larger than 10^{30}.

[35] Black (2012).

rating blanks), denoted for simplicity where each word is a capital letter. The number of possible combinations is 2N - 1 with N=number of name-words in the name and here N=3 with 7 combinations: (1) ABC, (2) A B, (3) A C, (4) B C, (5) A, (6) B, (7) C.

The ordering of the name-words does not matter because the function (using the name-word numeralizations) ensures an identical solution for any word order; the function is a multiplicative product of the numeralizations and multiplicative products are transitive. For instance, if the numeralizations of the name-words are (with simulated numerical values): A=10, B=27 and C=39, then function value equals the multiplication product of the three name-word's numerical values A*B*C=10530. Further, the product is the same regardless of the name-word order:

$$A^*B^*C = A^*C^*B = B^*A^*C = B^*C^*A = C^*A^*B = C^*B^*A = 10530,$$

for every possible ordering of the three-word name. This numeralization feature of name-word order equivalence greatly reduces the computational problem, compared to text methods, such as regular expressions or boulean equivalence.[36] In addition, transforming the matching into a numerical domain vastly reduces the memory required by each processing instruction since each digit in the numerical value of the solution of the comparator function can attain 10 values, while a single character in text has 256 potential values (EPCDIC) for standard character sets and 512 or more for extended character sets, such as those using special characters found in foreign languages used in the information reporting data record reported by the source information reporting country's tax administration.

The resulting set of PELname candidates for each recipient name is then ranked to select the "best" PELname candidate for each recipient. A PELname candidate detected in the first step of the matching process has one or more name-words that are common with the name-words in the name from the information reporting data record. The vector space model is used to quantify an index similar to what is used in Internet search engines.[37] The index reflects both the similarity and the dissimilarity of the PELname candidate and the recipient name. In general, the index is inversely correlated with the frequency of use of the name-words. Therefore, matching a word like "corporation", which is very common in the information reporting data names and PELnames, has far less consequence than for an uncommon word in an affiliate or entity name like Arapahoe or Stryker.

The resulting "best" PELname candidate for each recipient is further assigned a certainty measure that provides a statistical inference about the meaningfulness of the "best" match. In other words, not all "best" matches are meaningful. Assume a recipient entity or affiliate (a PELname) of a U.S. consolidated corporate taxpayer that earns income in a source information reporting country is properly identified as such (i.e., a U.S. corporate recipient) to the tax administration in that information reporting country. If so, then that recipient entity likely will be reported in the corporate tax return filed by the U.S. consolidated corporate taxpayer. The distribution of all PELname candidates, using the vector score of the PELname candidates as a random variable, is used to define the certainty of the match. Certainty is defined as those PELname candidates with vector scores that are more than two standard deviations above the mean (upper one-tail test statistics).

B. Payor Matching Methods

The payor names are also matched to PEL names to identify intercompany transactions. PELname candidates for each payor in the information reporting data are first derived using a two-step process similar to that used to match the recipients. However, the PELname list is restricted to entities and affiliates reported on three forms: Form 5471 for CFCs, Form 8858 for disregarded entities (DEs), and Form 8865 for foreign corporate partnerships (FCPs). All other PELnames are excluded from the list of potential PELname candidates in the

[36] The number of name-word orderings is N! where N=number of name-words in the name reported in the information reporting data or the number of name-words in the PELname. Since the average number of name-words is about 5 in a name, the numeralization and computational function reduces the number of comparator operations by a factor of about 1/N!=1/120—not to mention the efficacy of the computer memory required for the numeralizations instead of text character comparators.

[37] Salton, *et al.* (1975).

matching process. In addition, the number of unique payor names is much smaller than the number of unique recipient names because each payor may report many recipients. However, the average number of name-words of payor names is much larger than the average recipient name. Consequently, the size of the combinatorial problem for payors is actually larger than the combinatorial universe for recipients.

Once the PELname candidates for each FTIR payor have been selected, an additional processing step is applied to determine the "best" PELname payor candidate that is a meaningful match, using information about associated entities of the payors.[38] We expect that the payors and at least some of the recipients of that payor will be affiliated with the same consolidated corporate tax group reported in the U.S. tax return.

This condition of the associated entities suggests that of all the payor PELname candidates found for a payor name, the "best" match that is meaningful is that payor PELname candidate with a group EIN that also matches one or more of the group EINs for the matched PELname for the corresponding recipient. In other words, a related payor will have recipients that are also part of the same consolidated corporate tax group, and the payor and recipients have the same consolidated group EIN reported in the corporate tax return database that is found in the matching process.

Each FTIR record reports both the payor and the recipient. The PELname candidate for the payor includes the group EIN, and that is then compared to each of the EINS of the (previously) best matched PELname recipients. The "best" payor PELname candidate has a group EIN (of the consolidated tax group) that is common to one or more of the "best" match recipient PELnames. This condition demonstrates the existence of intra-company financial transactions between affiliated entities within the consolidated corporate tax group. When the matched EINs are found, this is deemed to identify the "best" PELname match that is meaningful for the payor name, and we conclude that this payor/recipient record is an intercompany transaction.

[38] The distribution of vector scores (VS) for payors candidates lacks properties of normality that would allow drawing conventional statistical inferences about meaningfulness. The recipient VS distribution has normality and 95%+ certainty can be associated with the properties of the normal distribution. However, because most payors do not have U.S. tax nexus, the payor candidate VS distribution would not be expected to have the same qualities.

Appendix C
Income-Shifting Variables

K_d Domestic capital stock is measured as the sum of buildings and other depreciable assets, depletable assets, land, intangibles, and other assets as reported on Form 1120, Schedule L, Lines 10 through 14. K_d must be greater than zero to enable calculation of return on assets.

K_f Foreign capital stock is measured as the sum of buildings and other depreciable assets, depletable assets, land, intangibles, and other assets as reported on Forms 5471, Schedule F, Lines 8 through 12. K_f must be greater than zero in all foreign jurisdictions to enable calculation of return on assets.

Y_d Domestic pretax income is measured as taxable income reported on Form 1120, Line 30 (TXBL_INCOM), less foreign-source dividends reported on Form 1120, Schedule C (Line 13: SCHC_OTH_FRGN_DIV, Line 14: SCHC_CNTRL_FRGN_INCM, Line 15: SCHC_FRGN_DIV_GR_UP).

Y_f Foreign pretax income is measured as the current earnings and profits reported on Forms 5471, Schedule H, Line 5(d), plus foreign tax reported on Forms 5471, Schedule E, Line 8(d).

T The tax liability associated with an entity's taxable income reported in an individual jurisdiction. The domestic tax variable, Td, is equal to the sum of total tax reported on Form 1120, line 31, plus the foreign tax credit reported on Form 1120, Schedule J, Line 6(a), multiplied by the ratio of domestic pretax income, Y_d, to taxable income reported on Form 1120, Line 30.

T_f The foreign tax variable, T_f, is equal to the sum of total income, war profits and excess profits taxes paid or accrued as reported on Forms 5471, Schedule E, Line 8(d).

ETR td; tf The effective tax rate, ETR, is calculated as T/Y by taxpayer by jurisdiction by year. Prior to calculating the income-shifting estimates, we truncate t at the EIN-level as follows: when $t > 1$, then t is set to 1; when $t < 0$, then t is set 0. The domestic tax rate is computed as domestic tax divided by domestic pretax income, or $ETR_d = T_d / Y_d$. The foreign tax rate is computed as foreign tax divided by foreign pretax income, or $ETR_f = T_f / Y_f$.

CNTRY Two-character identifier used to indicate the foreign jurisdiction associated with each filed Form 5471. CNTRY is defined as the principal place of business (PPB) reported in Box 1(e) on Form 5471, or if PPB is missing, CNTRY is the country of incorporation (COI) reported in Box 1(c) when known.

fyear We define TXY as the first four digits of TX_PRD_END_DR, an 8 digit variable, e.g., 20121231, FYR is the next two digits, the ending month of the fiscal year. We then define FYEAR as TXY unless FYR <6; then FYEAR = TXY -1. Thus, FYEAR agrees with Compustat's FYEAR.

The Economic Effects of Special Purpose Entities on Corporate Tax Avoidance

Paul Demeré (University of Georgia), and Michael P. Donohoe and Petro Lisowsky (University of Illinois at Urbana-Champaign)[1]

Both academic research and government reports study the tax effects of many different tax-advantaged *transactions*.[2] However, few studies evaluate the *organizational structures* facilitating such transactions, while those that do examine tax incentives and general determinants rather than the tax effects (Shevlin (1987); Beatty, *et al.* (1995); Feng, *et al.* (2009); Donohoe, *et al.* (2013)). Empirical evidence on the tax effects of organizational structures is important because, while lawmakers and regulators call for changes in the tax and financial reporting policies for organizational structures (JCT (2003); Basel Committee on Banking Supervision (2009); OECD (2013, 2015a, 2015b)), existing knowledge about the tax revenue losses attributable to such structures is largely anecdotal. We fill this void by examining whether, how, and the extent to which increasingly common components of organizational structures—special purpose entities (SPEs)—facilitate corporate tax avoidance.[3]

SPEs are separate legal entities created by a sponsor-firm to perform narrow, predefined business activities or series of transactions (Feng, *et al.* (2009)).[4] Corporate use of SPEs is large and growing, with nearly a quarter of all Compustat firms and one-half of S&P 500 firms using at least one SPE (Zion and Carcache (2003)). The number of SPE users in our sample has also increased by more than 600 percent from 1997 to 2011. While SPEs are used in common financial arrangements, such as leases and securitizations, anecdotes suggest companies use them to facilitate corporate tax avoidance; that is, reduce explicit taxes (JCT (2003)). To this end, critics consider special purpose entities a "series of dirty words" in taxation (Forbes and Sharma (2008)), while many other groups, including researchers (Mills, *et al.* (2012); Zion and Carcache (2003)), global tax authorities (Internal Revenue Bulletin 2011-39; Inland Revenue (2013)), and regulators (FASB (2003); JCT (2003, 2011); OECD (2013, 2015a, 2015b); United Nations (2013)) suspect that SPEs contribute to the continuing decline in corporate tax revenues.

Unlike tax-advantaged transactions, SPEs are organizational structures that do not directly generate tax savings, but instead facilitate tax savings in two ways.[5] The first way is by allowing sponsors to conduct a greater *level* of tax-advantaged transactions. By separating high-risk assets from the sponsor, SPEs can enable greater

[1] We appreciate helpful comments from participants at the American Taxation Association Midyear Meeting, the AAA Annual Meeting, the TPC-IRS Research Conference, National Tax Association Annual Conference, the Berlin-Vallendar Conference on Tax Research, and the PhD seminar at the EIASM/University of Münster Conference on Current Taxation. We also thank workshop participants at Arizona State University, Fordham University, the University of Connecticut, the University of Illinois at Urbana-Champaign, Vienna University of Economics and Business, as well as Andrew Bauer, Jenny Brown, Hye Sun Chang, Raluca Chiorean, Keith Czerney, Will Demeré, Katharine Drake, Brian Gale, Danielle Green, Jost Heckemeyer, Ryan Huston, Laura Li, Jeremy Lill, Sean McGuire, Tom Neubig, Tom Omer, George Plesko, Liora Schulman, Eric Toder, and Jaron Wilde. Special thanks to Richard Crowley for data collection assistance, and Mei Feng, Jeff Gramlich, and Sanjay Gupta for providing special purpose entity data. Paul Demeré gratefully acknowledges financial support from the AICPA Accounting Doctoral Scholars Program. Michael Donohoe gratefully acknowledges financial support from the PricewaterhouseCoopers Faculty Fellowship. Petro Lisowsky appreciates financial support from the MIT Sloan School of Management, as well as the PricewaterhouseCoopers Faculty Fellowship and Professor Ken Perry Fellowship at the University of Illinois at Urbana-Champaign.

[2] See U.S. Treasury (1999); Graham (2000); Hanlon and Heitzman (2010); Government Accountability Office [GAO] (2011); Organisation for Economic Co-Operation and Development [OECD] (2011); and U.S. Congress Joint Committee on Taxation [JCT] (2011).

[3] Tax avoidance does not necessarily imply improper behavior, as managing tax costs is an appropriate component of a firm's long-term strategy (Atwood, et al. (2012)). Our full working paper (available upon request) describes how SPEs facilitate tax avoidance.

[4] SPEs are also known as variable interest entities (VIEs) or special purpose vehicles (SPVs). VIEs are a subset of SPEs subject to consolidation under Financial Interpretation No. 46 (Financial Accounting Standards Board [FASB] (2006)) and subsequent pronouncements (Chasteen (2005)). We use the term "SPE" to refer to all such entities.

[5] A tax-advantaged transaction reduces tax costs independent of taxpayer motives. For example, the debt-tax shield of leverage and tax credits from R&D expense reduce taxes, but do not necessarily result from intentional tax planning. We discuss this concept further and map empirical proxies to such transactions in our full working paper available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2557752.

(1) debt capacity, resulting in more tax-deductible interest expense (Mills and Newberry (2005)); (2) external financing, leading to more research and development (R&D) deductions and tax credits (Shevlin (1987)); and (3) synthetic leases, increasing depreciation deductions (Zechman (2010)). The second way that SPEs facilitate tax savings is by enhancing tax *efficiency*, or the relative tax savings from existing tax-advantaged transactions (i.e., holding *level* constant). For example, SPEs allow firms to: (1) shift profits to jurisdictions with low tax rates (Drucker (2007); Forbes and Sharma (2008); Dyreng, *et al.* (2013)); (2) design and operate tax shelters (Graham and Tucker (2006); Wilson (2009); Lisowsky (2010)); and (3) structure intercompany transactions that result in tax credit and loss duplication (JCT (2003); Sheppard (2017)).

It is important to understand the ways in which SPEs facilitate tax-advantaged transactions (i.e., by changing their level or efficiency) because they shed light on the tax-motivated business strategies to achieve tax savings (Scholes, *et al.* (2014)). For example, if SPEs enable a firm to engage in a greater level of tax-advantaged transactions, such as R&D, then this tax avoidance may be within the bounds of tax law and beneficial to corporate stakeholders. However, if SPEs enable a firm to enhance the tax efficiency of a transaction, such as shifting R&D-related profits to tax havens or implementing a tax shelter that results in R&D credit duplication, then this tax avoidance may be pushing the bounds of tax law and exposing corporate stakeholders to additional costs (e.g., tax audits and penalties).

Despite our focus on taxes, some experts argue that obtaining tax savings is *not* the main objective of common SPEs (e.g., for asset financing), suggesting these organizational structures play a minor role in corporate tax avoidance (Soroosh and Ciesielski (2004)). Even when tax savings are a primary objective, SPEs are not necessarily optimal once all costs are considered (Scholes, *et al.* (2014)). For example, in addition to legal setup costs, SPEs can reduce information quality (Feng, *et al.* (2009)), increase regulatory scrutiny (Internal Revenue Bulletin 2011–39; Inland Revenue (2013)), enhance public pressure (Dyreng, *et al.* (2016)), and result in large tax penalties (Wilson (2009)).

Accordingly, we investigate whether and under what circumstances the tax effects of SPEs are economically significant by answering three open empirical questions. First, to what extent do SPEs enable sponsor-firms to conduct a greater *level* of tax-advantaged transactions? That is, we examine which specific transactions are commonly used within SPE-structures for tax avoidance. Second, how large are the tax savings facilitated by SPEs? Third, for which specific transactions do SPEs enhance tax *efficiency*, or relative tax savings? Directly measuring the level and efficiency of tax savings facilitated by SPEs will empirically answer important tax policy questions; namely, whether, how, and the extent to which organizational structures enable corporate tax avoidance.

We begin our analyses by providing the first large-sample empirical evidence on the overall relation between SPEs and corporate tax avoidance. Specifically, we regress two different forward-looking effective tax rates (ETRs) estimated over a three-year horizon (t to t+2) on both a binary and continuous measure of SPEs derived from Exhibit 21 of Form 10-K during 1997–2011. We use fixed-effects estimation (a generalized difference-in-differences framework), which measures the effect of changes in SPE use on ETRs; that is, the *incremental* tax savings attributable to SPEs (Wooldridge (2010); Roberts and Whited (2013)). We find that both the number of SPEs and the use of SPEs are negatively and significantly associated with future GAAP ETRs (i.e., total tax expense scaled by pretax income) and cash ETRs (i.e., cash taxes paid scaled by pretax income), suggesting that SPEs facilitate tax avoidance incremental to common tax-advantaged transactions and other controls.

Next, we use path analysis to decompose the overall relation between SPEs and ETRs into direct and indirect paths. While regression analysis gauges overall effects, path analysis considers the existence and relative importance of alternative (indirect) paths of influence that jointly create overall effects (Bhattacharya, *et al.* (2012)). As a class of structural equation models, path analysis allows us to investigate the extent to which tax-advantaged transactions captured in our model are used *within* SPEs to avoid taxes. In other words, path analysis estimates the extent to which SPEs incrementally increase the *level* of tax-advantaged transactions, and provides a focused setting in which to estimate the total tax savings facilitated by SPEs.

We find that several tax-advantaged transactions are used within SPEs to avoid corporate taxes. Specifically, we estimate that SPEs facilitate 1.8 percent of the cash tax savings from leverage, 3.3 percent from net

operating loss (NOL) carryforwards, 8.7 percent from R&D, 6.1 percent from intangibles, and all of the cash tax savings from tax havens. Thus, the path analysis indicates that SPEs increase the *level* of several specific tax-advantaged transactions in an economically meaningful way.

The path analysis also reveals that SPEs facilitate an economically significant amount of total tax savings. Among our principal results, we find that firms using the mean number of SPEs (5.48) have cash ETRs that are 4.4 percent lower than nonusers. At the firm level, these effects indicate that SPE users realize $7.8 million more in cash tax savings per year than firms not using SPEs. In aggregate, we estimate total cash tax savings of $82.4 billion for our sample of 10,284 SPE users, or approximately 2 percent of total U.S. federal corporate income tax collections during the sample period.[6] These estimates are considerably larger than those for other complex planning strategies, including tax shelters (Wilson (2009)), tax havens (Dyreng and Lindsey (2009)), round-tripping (Hanlon, *et al.* (2015)), and financial derivatives (Donohoe (2015)).

We next use moderation analysis to estimate the extent to which SPEs enhance the relative tax savings (i.e., tax *efficiency*) of tax-advantaged transactions. Moderation analysis considers if the relation between two variables depends on a third variable, allowing us to examine if specific transactions (as captured by model covariates) generate more or less tax savings when performed within versus outside SPEs. While the *path* analysis reveals that an economically large portion of the tax savings from leverage, NOLs carryforwards, and tax havens occur within SPEs by contributing to increased debt capacity, loss deductibility, and income shifting opportunities, respectively, the *moderation* analysis shows that SPEs do not enhance the tax efficiency of these transactions. However, SPEs enable a greater level *and* efficiency of total tax savings for R&D and intangibles-based transactions by 92.6 percent and 72.5 percent, respectively.

Finally, we perform several other tests to provide further insight on the tax effects of SPEs. First, we consider the link between SPEs and tax aggressiveness. We find that SPE use has a positive relation with unrecognized tax benefits (Lisowsky, *et al.* (2013)), but not the likelihood of tax shelter participation (Lisowsky (2010)), suggesting that, on average, SPEs facilitate some tax uncertainty, but not extremely aggressive positions.[7] Second, we find that the GAAP ETR results are stronger for U.S. multinationals compared to U.S. domestic firms, and a majority of the tax savings of SPEs comes from avoiding U.S. Federal, rather than foreign or State, income taxes. Our results also hold across several industries, suggesting that SPE-facilitated tax avoidance is pervasive and not simply confined to high-tech or intangible-intensive firms. Third, we mitigate alternative explanations by showing that our results are not driven by: (1) the endogenous choice to use SPEs; (2) the financial reporting of minority owners of SPEs; (3) variation in firms' subsidiary reporting over time; or (4) potential increases in overall organizational complexity.

This study contributes to the literatures on SPEs (Shevlin (1987); Beatty, *et al.* (1995); Feng, *et al.* (2009)) and corporate tax avoidance (see Hanlon and Heitzman (2010)) in three ways. First, we differ from traditional tax avoidance research in that we consider whether, how, and the extent to which tax avoidance is facilitated by increasingly common and uniquely complex organizational *structures*. In doing so, we identify some of the transactions used within such structures to facilitate tax savings, which is relevant to market participants as they analyze firms' tax profiles (Weber (2009)), and to tax authorities as they evaluate enforcement efforts to combat declining corporate tax revenues (Fox and Luna (2005); Inland Revenue (2013); Bozanic, *et al.* (2017); Dyreng, *et al.* (2017)). Second, by providing the first large-sample estimates of the total tax savings facilitated by SPEs, we clarify inconclusive anecdotal evidence routinely cited by experts (e.g., Zion and Carcache (2003); Soroosh and Ciesielski (2004); Forbes and Sharma (2008)) and researchers (e.g., Chasteen (2005); Desai and Dharmapala (2006); Feng, *et al.* (2009); Zechman (2010)), as well as help move the literature beyond the notion that firms simply *can* use SPEs to facilitate tax savings. Further, while prior research finds that tax incentives are an important—but not leading—determinant of SPE use, we show that SPEs facilitate economically significant tax savings nonetheless. Finally, our study is the first to use both path and moderation analysis to: (1) evaluate the tax effects of organizational structure; and (2) separate level from efficiency effects. It can thus guide future research examining the economic outcomes of other corporate organizational structures.

[6] U.S. Federal corporate income tax collections totaled approximately $4.46 trillion from 1997 to 2013 (https://www.irs.gov/statistics/soi-tax-stats-collections-and-refunds-by-type-of-tax-irs-data-book-table-1). While our sample spans 1997–2011, our tests use data through 2013 to calculate forward-looking effective tax rate measures.

[7] Tax aggressiveness is typically considered the use of tax positions that "push the envelope of tax law" (Hanlon and Heitzman (2010)) and is a subset of tax avoidance (Lisowsky, *et al.* (2013)).

References

Atwood, T.J., M. Drake, J. Myers, and L. Myers. 2012. Home Country Tax System Characteristics and Corporate Tax Avoidance: International Evidence. *The Accounting Review* 87(6): 1831–1860.

Basel Committee on Banking Supervision (Basel). 2009. *Report on Special Purpose Entities*. Basel, Switzerland: Bank for International Settlements.

Beatty, A., P. Berger, and J. Magliolo. 1995. Motives for Forming Research & Development Financing Organizations. *Journal of Accounting and Economics* 19: 411–442.

Bhattacharya, N., F. Ecker, P. Olsson, and K. Schipper. 2012. Direct and Mediated Associations Among Earnings Quality, Information Asymmetry, and the Cost of Equity. *The Accounting Review* 87(2): 449–482.

Bozanic, Z., J. Hoopes, J. Thornock, and B. Williams. 2017. IRS Attention. *Journal of Accounting Research* 55 (1): 79–114.

Chasteen, L. 2005. Teaching Variable Interest Entities Under FIN 46: Untangling Risks, Expected Losses, and Expected Residual Returns. *Journal of Accounting Education* 23: 47–66.

Desai, M., and D. Dharmapala. 2006. Corporate Tax Avoidance and High-Powered Incentives. *Journal of Financial Economics* 79: 145–179.

Donohoe, M., G. McGill, and E. Outslay. 2013. Back to the Drawing Board: The Structural and Accounting Consequences of a Switch to a Territorial Tax System. *National Tax Journal* 66(3): 713–744.

Donohoe, M. 2015. The Economic Effects of Financial Derivatives on Corporate Tax Avoidance. *Journal of Accounting and Economics* 59: 1–24.

Drucker, J. 2007. Wal-Mart Cuts Taxes by Paying Rent to Itself. *Wall Street Journal,* February 1.

Dyreng, S., M. Hanlon, E. Maydew, and J. Thornock. 2017. Changes in Corporate Effective Tax Rates Over the Past 25 Years. *Journal of Financial Economics* 124: 441–463.

Dyreng, S., J. Hoopes, and J. Wilde. 2016. Public Pressure and Corporate Tax Behavior. *Journal of Accounting Research* 54(1): 147–185.

Dyreng, S., and B. Lindsey. 2009. Using Financial Accounting Data To Examine the Effect of Foreign Operations Located in Tax Havens and Other Countries on U.S. Multinational Firms' Tax Rates. *Journal of Accounting Research* 47(5): 1283–1316.

Dyreng, S., B. Lindsey, and J. Thornock. 2013. Exploring the Role Delaware Plays as a Tax Haven. *Journal of Financial Economics* 108: 751–772.

Feng, M., J. Gramlich, and S. Gupta. 2009. Special Purpose Vehicles: Empirical Evidence on Determinants and Earnings Management. *The Accounting Review* 84(6): 1833–1876.

Financial Accounting Standards Board (FASB). 2003. *Consolidation of Variable Interest Entities—An Interpretation of ARB No. 51*. FASB Interpretation No. 46 (revised). Norwalk, CT: FASB.

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes—An Interpretation of FASB Statement No. 109*. FASB Interpretation No. 48. Norwalk, CT: FASB.

Financial Accounting Standards Board (FASB). 2009. *Amendments to FASB Interpretation No. 46(R)*. Statement of Financial Accounting Standards No. 167. Norwalk, CT: FASB.

Forbes, T., and R. Sharma. 2008. Getting Tax Value from Intangibles and SPVs. *International Tax Review* 19(9): 16–17.

Fox, W., and L. Luna. 2005. Do Limited Liability Companies Explain Declining State Corporate Tax Revenues? *Public Finance Review* 33(6): 690–720.

Government Accountability Office (GAO). 2011. Financial Derivatives: Disparate Tax Treatment and Information Gaps Create Uncertainty and Potential Abuse. Report to Congressional Requesters. September.

Graham, J. 2000. How Big Are the Tax Benefits of Debt? *The Journal of Finance* 55(5): 1901–1941.

Graham, J., and A. Tucker. 2006. Tax Shelters and Corporate Debt Policy. *Journal of Financial Economics* 81: 563–594.

Hanlon, M., and S. Heitzman. 2010. A Review of Tax Research. *Journal of Accounting and Economics* 50: 127–178.

Hanlon, M., E. Maydew, and J. Thornock. 2015. Taking the Long Way Home: U.S. Tax Evasion and Offshore Investments in U.S. Equity and Debt Markets. *Journal of Finance* 70(1): 257–287.

Inland Revenue. 2013. Multinational Enterprises: Compliance Focus (IR 947). New Zealand Government Inland Revenue. October.

Internal Revenue Bulletin 2011-39. Available at: https://www.irs.gov/pub/irs-irbs/irb11-39.pdf.

Lisowsky, P. 2010. Seeking Shelter: Empirically Modeling Tax Shelters Using Financial Statement Information. *The Accounting Review* 85(5): 1693–1720.

Lisowsky, P., L. Robinson, and A. Schmidt. 2013. Do Publicly Disclosed Tax Reserves Tell Us About Privately Disclosed Tax Shelter Activity? *Journal of Accounting Research* 51(3): 583–629.

Mills, L., and K. Newberry. 2005. Firms' Off-Balance Sheet and Hybrid Debt Financing: Evidence from Their Book-Tax Reporting Differences. *Journal of Accounting Research* 43(2): 251–282.

Organisation for Economic Co-operation and Development (OECD). 2011. Corporate Loss Utilization Through Aggressive Tax Planning. Paris, France: OECD Publishing.

Organisation for Economic Co-operation and Development (OECD). 2013. Revised Discussion Draft on Transfer Pricing Aspects of Intangibles. July 30.

Organisation for Economic Co-operation and Development (OECD). 2015a. Action 6: Preventing the Granting of Treaty Benefits in Inappropriate Circumstances. October 5, 2015.

Organisation for Economic Co-operation and Development (OECD). 2015b. Action 11: Measuring and Monitoring BEPS. October 5, 2015.

Roberts, M., and T. Whited. 2013. Endogeneity in Empirical Corporate Finance. *Handbook of the Economics of Finance* 2(A): 1–798.

Scholes, M., M. Wolfson, M. Erickson, M. Hanlon, E. Maydew, and T. Shevlin. 2014. *Taxes and Business Strategy: A Planning Approach*. Upper Saddle River, NJ: Pearson Education, Inc.

Sheppard, L. 2017. Partnership Basis Stripping Inside a Consolidated Return. *Tax Notes* (July 24): 385–394.

Shevlin, T. 1987. Taxes and Off-Balance-Sheet Financing: Research and Development Limited Partnerships. *The Accounting Review* 62(3): 480–509.

Soroosh, J., and J. Ciesielski. 2004. Accounting for Special Purpose Entities Revised: FASB Interpretation 46(R). *The CPA Journal* 74(7): 30–37.

United Nations. 2013. World Investment Report. United Nations Conference on Trade and Development. Geneva, Switzerland: United Nations.

U.S. Congress Joint Committee on Taxation (JCT). 2003. *Report of Investigation of Enron Corporation and Related Entities Regarding Federal Tax and Compensation Issues, and Policy Recommendations* (JCS-3-03). February.

U.S. Congress Joint Committee on Taxation (JCT). 2011. Present Law and Issues Related to the Taxation of Financial Instruments and Products (JCX-56-11). December 2.

U.S. Department of the Treasury. 1999. *The Problem of Corporate Tax Shelters: Discussion, Analysis and Legislative Proposals*. July. Washington, DC: Government Printing Office.

Weber, D. 2009. Do Analysts and Investors Fully Appreciate the Implications of Book-Tax Differences for Future Earnings? *Contemporary Accounting Research* 26(4): 1175–1206.

Wilson, R. 2009. An Examination of Corporate Tax Shelter Participants. *The Accounting Review* 84(3): 969–999.

Wooldridge, J. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second Edition. Cambridge, MA: The MIT Press.

Zechman, S. 2010. The Relation Between Voluntary Disclosure and Financial Reporting: Evidence from Synthetic Leases. *Journal of Accounting Research* 48(3): 725–765.

Zion, D., and B. Carcache. 2003. FIN 46: New Rule Could Surprise Investors. Credit Suisse/First Boston Equity Research. June 24.

2

Realizing the Potential of Tax Enforcement

Michelle Nessa ◆ Casey Schwab ◆ Bridget Stomberg
Erin Towery

Elena D'Agosto ◆ Marco Manzo ◆ Alessandro Modica
Stefano Pisani

Alex Turk ◆ Eric Henry ◆ Dan Howar
Maryamm Muzikir

How Do IRS Resources Affect the Tax Enforcement Process?

Michelle Nessa (Michigan State University), Casey Schwab and Bridget Stomberg (Indiana University), and Erin Towery (University of Georgia) [1]

Recent cuts to the Internal Revenue Service's (IRS) budget have reduced its resources by 18 percent and resulted in 13,000 (14 percent) fewer employees, 10,000 (20 percent) fewer enforcement staff, and the lowest level of individual and business audits in a decade (Marr and Murray (2016)). These cuts have occurred despite a 4-percent increase in the number of returns filed since 2010 and other increased IRS responsibilities related to the implementation of the Foreign Account Tax Compliance Act and the Affordable Care Act. Numerous media articles provide anecdotal evidence of the negative repercussions that have already occurred and speculate about potential future consequences (e.g., Rubin (2015); Russell (2015)). In this study, we use confidential IRS data to examine the effect of IRS resources on each stage of the enforcement process including: (1) the rate of audit; (2) the incidence and magnitude of proposed deficiencies conditional on audit; and (3) the percentage of proposed deficiencies collected by the IRS. In doing so, we provide the first large-sample evidence on how IRS resources affect each stage of the enforcement process—particularly among large corporations.

While prior research documents a negative association between audit probabilities and corporate tax avoidance (Hoopes, Mescall, and Pittman (2012)), there is limited evidence regarding how IRS resources affect enforcement outcomes. Although conventional wisdom might suggest greater resources are associated with better enforcement outcomes, this need not be the case for at least two reasons. First, over the last 15 years, the IRS has enhanced the tools it uses to detect potential noncompliance. For example, a joint IRS and Department of Treasury work group created the Schedule M-3 in 2003 to require greater disaggregation of firms' book-tax differences. The increased quantity and standardization of book-tax difference disclosures was intended to enable the IRS to better analyze book-tax differences for compliance risks, while simultaneously reducing the likelihood that the IRS will pursue a return based on an incorrect assumption about aggressive tax reporting (Boynton and Mills (2004)). Other initiatives include the creation and implementation of: (1) the Modernized e-File system and mandatory electronic tax return filing for certain corporations; (2) the Compliance Assurance Process (CAP), which facilitates resolution of questionable issues prior to the return filing; (3) Forms 8886 and 8918, which require a taxpayer and material advisors to the taxpayer to disclose reportable transactions; and (4) Schedule UTP, which requires additional disclosures related to a taxpayer's uncertain tax positions. These initiatives are intended to reduce the time and costs of identifying and resolving uncertain and/or more aggressive tax positions and could enable the IRS to maintain the level of tax collections despite resource reductions.[2] Second, House Appropriations Committee Chairman Rogers suggested that the rationale for recent IRS budget cuts was an effort to streamline "inefficient" operations (Bedard (2014)). This characterization reflects the belief that lower IRS resources will not hamper—and could potentially enhance—the enforcement process.

We focus our analysis on how IRS resources affect IRS enforcement after the tax return is filed. The IRS faces a number of decisions when allocating enforcement resources, including the number of returns to audit,

[1] The authors appreciate helpful comments from Terrence Blackburne, Lisa De Simone, Susanna Gallani, Joanna Garcia, Ken Klassen, Pete Lisowsky (discussant), Lillian Mills, Kathleen Powers, Phil Quinn, Sonja Rego, Jeri Seidman, Jake Thornock, Alex Turk (discussant), Michael Udell (discussant), Brian Williams, Terry Yohn, and workshop participants at the University of Washington, the University of Waterloo, Washington State University, University of Houston, North Carolina State University, Indiana University, Singapore Management University, the 2016 *Journal of the American Taxation Association* Conference, the 2016 National Tax Association Annual Meeting, and the 2017 IRS-TPC Research Conference. We thank Scott Dyreng (Duke University) for providing data on the location of firms' foreign subsidiaries. We also thank John Miller and Barbara Hecimovich for providing information about the IRS audit process.

[2] The Treasury Inspector General for Tax Administration (TIGTA) concluded that the Modernized e-File system and mandatory electronic filing: (1) reduced the time and costs associated with processing tax returns; (2) decreased the amount of time needed to complete audits; and (3) increased deficiencies proposed during audits (TIGTA (2011)).

the scope of the audits (which influences the incidence and magnitude of proposed deficiencies), and the level of resources dedicated to collecting proposed deficiencies. Because the IRS' budget is fixed for any fiscal year, resource allocation is a zero sum game: resources consumed in early stages of the enforcement process affect resources available for later stages. As such, it is unclear how the level of IRS resources affects each stage of the enforcement process or whether any changes in collections stem from agents examining a different number of returns and/or altering the resources devoted to uncovering, investigating, and challenging potential noncompliance.

To test the effect of IRS resources on enforcement, we use confidential corporate audit examination data for audits conducted from 2002 through 2014 of tax return years from 2000 through 2010.[3] To construct our sample, we use employer identification numbers to merge public financial statement data from the Compustat Fundamentals Annual database with three IRS datasets: (1) the IRS Business Return Transaction File (BRTF), which contains corporate income tax return data; (2) the IRS Audit Information Management System (AIMS), which contains proposed tax deficiencies; and (3) the IRS Enforcement Revenue Information System (ERIS), which contains deficiencies collected by the IRS. We restrict the sample to publicly traded firms so that we can include a more comprehensive set of control variables than we could if the sample included both public and private firms. We obtain data on IRS resources from the *IRS Annual Data Book*. Our primary measures of IRS resources are the IRS' total enforcement budget and the number of revenue agents. When examining the probability of IRS audit, we scale each measure by the number of returns filed. When examining later stages in the enforcement process that are contingent on a firm being audited, we scale each measure by the number of returns audited.

We first provide descriptive information about our sample. Approximately 29 percent of return years not in the Coordinated Industry Case (CIC) program are audited by the IRS. For the sample of audited returns, roughly 49 percent receive a proposed deficiency from the IRS, and the average magnitude of the proposed deficiency is almost $5.2 million, which represents approximately 6.8 percent of the tax savings claimed on the originally filed return. On average, corporate taxpayers pay 73 percent of proposed deficiencies—most of the shortfall due to proposed deficiencies not being assessed as payable.

We conduct our primary analyses in three stages. First, we examine the impact of IRS resources on the probability of audit among corporations. Using a sample of 31,549 tax return-years, we find a positive association between IRS resources and the probability of audit. A one standard deviation reduction in the IRS enforcement budget is associated with a 2.3-percent reduction in the probability of audit relative to the base rate. Second, we examine the impact of IRS resources on the incidence and magnitude of proposed deficiencies. We find a positive association between IRS resources and both the incidence and magnitude of proposed deficiencies within a sample of 11,899 audited corporation tax return-years. A one standard deviation reduction in the IRS enforcement budget is associated with an 11.1-percent decrease in the magnitude of proposed deficiencies relative to the mean level of proposed deficiencies. These results are consistent with the prospect that lower levels of IRS resources weaken the initial stages of the enforcement process, with the IRS auditing fewer returns and proposing fewer and smaller deficiencies.

Finally, we examine the impact of IRS resources on negotiated settlement outcomes. We measure settlement outcomes as the proportion of proposed deficiencies collected by the IRS (the settlement ratio). A higher (lower) settlement ratio indicates better outcomes for the IRS (taxpayer). Using a sample of 5,840 audited tax return-years with a proposed deficiency, we find a negative association between IRS resources and settlement ratios. A one standard deviation reduction in the IRS enforcement budget is associated with a 1.7 percentage point increase in the proportion of proposed deficiencies retained by the IRS among our sample of corporation return-years with proposed deficiencies. This result is consistent with the IRS targeting weaker taxpayer positions when it has fewer resources. We further decompose total settlements into settlements collected following the initial examination and settlements collected following an appeal. Our results suggest that the IRS collects a larger portion of proposed deficiencies after the initial examination stage when it is has fewer resources. We find no effect of IRS resources on collections after appeals.

[3] We limit the sample to returns through the 2010 tax year to allow sufficient time for tax return examinations to be completed.

This study is important to Congress, tax authorities, and taxpayers because our results inform the ongoing political debate about IRS funding. Although we provide evidence that settlement ratios are higher when the IRS has fewer resources, we also report a reduction in the rate of audit and in the incidence and magnitude of proposed deficiencies. Our estimates collectively suggest that a reduction in IRS resources is associated with a net loss in revenue collected via the corporation tax enforcement process. Taxpayers should also be aware of IRS resources during the tax enforcement process because the IRS proposes fewer adjustments when it has fewer resources but sustains a larger proportion of those proposed adjustments.

Our study also contributes to the stream of academic research examining the interaction between corporations and the tax authority. Studies in this area often focus on taxpayer strategy in deciding which positions to claim (e.g., Mills, Robinson, and Sansing (2010); De Simone, Sansing, and Seidman (2013); Ayers, Seidman, and Towery (2017)) or on the determinants of proposed deficiencies (Mills (1998); Mills and Sansing (2000)). Nearly all models of taxpayer-tax authority interaction assume that noncompliance detected by the tax authority necessarily results in additional tax payments (Slemrod, Blumenthal, and Christian (2001) is a notable exception). In contrast, we conduct a comprehensive analysis of how IRS resources affect each stage in the enforcement process, including the outcomes of negotiations between corporations and the IRS *after* noncompliance is alleged. These negotiations are a significant aspect of the interaction between taxpayers and the tax authority and have important implications for government collections and taxpayer cash flows.

References

Ayers, B., J. Seidman, and E. Towery. 2017. Taxpayer behavior under audit certainty. Working paper, University of Georgia and University of Virginia.

Bedard, P. 2014. House budget punishes IRS with 15% cut, halts Obamacare enforcement. Available online at: http://www.washingtonexaminer.com/house-budget-punishes-irs-with-15-cut-halts-obamacare-enforcement/article/2549830. Last accessed August 29, 2016.

Boynton, C., and L. Mills. 2004. The evolving Schedule M-3: a new era of corporate show and tell? *National Tax Journal* 57(3): 757–772.

De Simone, L., R. Sansing, and J. Seidman. 2013. When are enhanced relationship tax compliance programs mutually beneficial? *The Accounting Review* 88(6): 1971–1991.

Hoopes, J., D. Mescall, and J. Pittman. 2012. Do IRS audits deter corporate tax avoidance? *The Accounting Review* 87(5): 1603–1639.

Marr, C., and C. Murray. 2016. IRS funding cuts compromise taxpayer service and weaken enforcement. Available online at: https://www.cbpp.org/research/federal-tax/irs-funding-cuts-compromise-taxpayer-service-and-weaken-enforcement. Last accessed August 29, 2016.

Mills, L. 1998. Book-tax differences and Internal Revenue Service adjustments. *Journal of Accounting Research* 36: 343–356.

Mills, L., L. Robinson, and R. Sansing. 2010. FIN 48 and tax compliance. *The Accounting Review* 85(5): 1721–1742.

Mills, L., and R. Sansing. 2000. Strategic tax and financial reporting decisions: theory and evidence. *Contemporary Accounting Research* 17(1): 85–106.

Rubin, R. 2015. IRS audits fall to an 11-year low. *Wall Street Journal*: November 4.

Russell, R. 2015. NTA highlights decline in taxpayer service. *Accounting Today*: February 1.

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: 455–483.

TIGTA. 2011. Successfully processing large corporate tax returns electronically was a major accomplishment, but eliminating more compliant returns from the audit stream is a work in progress. (Reference Number: 2011–30–048).

Tax Audits and Tax Compliance—Evidence from Italy

Elena D'Agosto, Marco Manzo, Alessandro Modica, and Stefano Pisani (Agenzia delle Entrate and Ministero dell'Economia e delle Finanze-Dipartimento delle Finanze, Government of Italy) [1]

1. Introduction

In Italy, tax evasion (including evasion of social security contributions) has been estimated to be 110 billion euros a year, about 7 percent of the GDP.[2] There exist a number of studies and reports showing that levels of compliance with tax laws are low in Italy compared to European standards.

In times of tightening fiscal constraints, combatting tax evasion and stimulating tax compliance have become a priority for economic policy. In this context, the most cost-effective strategy would be to use tax audit programs not only to directly detect tax evasion but also, and foremost, to promote voluntary compliance.

In this paper, we focus on the effect of tax audits on a sample of small firms and self-employed individuals over the period 2006-2011, using a unique and confidential dataset with tax return data and information on the tax audit activity of the Italian Revenue Agency (IRA). We investigate this particular group of firms for two reasons: first of all, among this group of taxpayers, the propensity to underreport income is higher than the national average; and second, the distribution of firms by size is skewed towards small and micro firms, which account for more than 98 percent of the total in Italy.[3]

Building on the results of a companion paper (D'Agosto, et al. (2017)), the aim of this study is to investigate whether tax audits exert a positive and relevant effect on tax compliance and to evaluate the effectiveness of different categories of audit programs. The final part of the paper is devoted to the analysis of the general deterrence effect of tax audits on tax compliance.

As most of the IRA audit programs are not random, we use an econometric approach that allows us to account for selection bias related to time-invariant individual characteristics. Moreover, our results seem robust to the endogeneity issues caused by the possible correlation between time-variant individual characteristics and the audit selection process.

The rest of the paper is organized as follows: the second section reviews the literature on tax compliance; the third section describes the dataset; section 4 describes the empirical model and the econometric approach used; and section 5 concludes.

2. Review of Related Literature

Since the pioneering Allingham and Sandmo (1972) study, economic research on taxpayer compliance has examined the role of enforcement to discourage fraudulent tax behavior. Empirically, studies have discerned two main effects of enforcement on tax collection. The first is the direct benefit from the additional tax, interest, and penalties paid by the audited taxpayer for the year(s) under audit. The second is an indirect effect, mainly defined in literature as a general deterrent effect, which is the change in tax collections among the general population—including those who were not audited.[4]

---

[1] Opinions expressed in this article are those of the authors and do not necessarily reflect those of the Finance Department and of the Revenue Agency. Any errors are ours.

[2] See "RELAZIONE SULL'ECONOMIA NON OSSERVATA E SULL'EVASIONE FISCALE E CONTRIBUTIVA" available here: http://www.mef.gov.it/inevidenza/documenti/Relazione_evasione_fiscale_e_contributiva__0926_ore1300_xversione_definitivax-29_settembre_2016.pdf.

[3] See http://www.istat.it/it/files/2015/05/CAP-3-Rapporto-Annuale-2015–2.pdf.

[4] Within the indirect impacts, some scholars define the effects spread upon nonaudited taxpayers as a spillover deterrent effect and identify the subsequent year effect of those audited as a corrective impact (Gemmell and Ratto, 2012). An empirical estimation of the enforcement spillovers is shown in Rincke and Traxler (2011).

The direct outcome of investigations may be measured immediately, the indirect outcome, however, is more difficult to estimate, as the contrasting results in literature have shown.

In this perspective, some studies have explored the effects of an increasing probability of an audit: for instance those of Dubin and Wilde (1988), Dubin, Graetz, and Wilde (1990), Plumley (1996) and Dubin (2007). These authors have used aggregated time series and cross-sectional data, data on individuals at the State level, and audit rate data from the IRS. In evaluating the direct and indirect effects on revenue, they showed that the indirect prevails over the direct. Recent studies, such as those of Birskyte (2013) and Tagkalakis (2014), have obtained similar results.

Another approach has looked at the effects that the experience of an audit has produced on taxpayer behavior by using individual data. In this setting, Erard (1992) examined the effect of an audit on subsequent year reporting behavior and his results, although consistent with a positive influence, are not conclusive. Niu (2011) investigates the voluntary compliance shift after a firm is audited by applying a difference-in-differences approach to tax return and audit data. His findings show that audited firms report higher sales growth rate in the year of audit than nonaudited firms do. Gemmell and Ratto (2012) analyze the subsequent compliance behavior of audited taxpayers and their results are inconclusive. Ratto*, et al.* (2013) try to formalize the direct and indirect effects of an audit from a theoretical point of view. Interestingly, the authors derive the expression of direct and indirect effect of audit on tax compliance by introducing the behavioral response in terms of elasticity of evasion. They show that their ratio depends on behavioral elasticity, on the intensity with which a group is controlled and on the operational effectiveness of investigations.

In Italy, few studies have analyzed the effects of enforcement policies. The main contributions are due to Santoro (2008), and Santoro and Fiorio (2011). Both are concerned with the effectiveness of Sector Studies (SdS), an audit tool adopted by the Italian Tax Revenue Agency to increase compliance. A further study by Fiorio*, et al.* (2013) shows the positive effect on revenue of the threat of an audit.[5] Di Porto (2011) analyzes the impact of tax inspections on employer labor tax evasion. His analysis is based on a data set of individual artisan firms and individual audit data. However, he finds that inspections may decrease both tax compliance and tax revenues.

Recently, D'Agosto*, et al.* (2017) explore the compliance behaviors of small businesses by means of unique datasets from the Italian Revenue Agency. The analysis focuses on the tax compliance effects of various enforcement policies: field audit (i.e., soft audit), desk audit (i.e., deep audit), and a combination of field and desk audit activities, over the period 2004–2009. The authors measure the impact on individual tax compliance in terms of changes in the tax declared by the audited taxpayers with respect to the nonaudited ones. Findings show that each policy has a positive and significant effect on compliance, although with different magnitudes.

3. The Dataset

We use a unique and confidential dataset, which gathers data from tax returns, with detailed information on the tax audit activity of the IRA. In particular, the dataset contains information on the VAT (Value Added Tax), the PIT (Personal Income Tax), and the IRAP (Regional Tax on Business Activities) tax return for a sample of small businesses over the period 2006–2011. Moreover, we are able to identify the taxpayers in the sample that were audited, the date when the tax audit occurred, and the type of tax audit.

With regards to the timing of the audits, we assume that audited taxpayers can change their behavior immediately after having been informed that an inspection is taking place on their fiscal standing. Our assumption is that the audit effect on tax compliance can be observed only if the taxpayers become aware of the tax inspection before having paid the first installment of the tax due.[6] As in D'Agosto*, et al.* (2017), we consider that an audit has taken place in year t only if the taxpayer receives the first notice of the inspection between the 1st of July of year t and the 30th of June of year t+1.

[5] See Blumenthal*, et al.* (2001) for an analogous experiment in Minnesota.

[6] Even though the tax calendar differs among VAT, PIT, and IRAP and between type taxpayers, setting the deadline for paying the first installment at the end of June is the best approximation as it includes the majority of cases.

Our dataset contains detailed information on different types of audit programs carried out by the IRA. In particular, we identify two broad categories of tax audits: "desk audits" and "on-site audits."

A desk audit is normally performed in the IRA offices and requires a deep investigation on all the budget figures of the taxpayers using all the available evidence (including the results of previous inspections). When there is evidence that an audited taxpayer understated or omitted real income, the IRA performs a desk audit to determine the true taxable income and tax of the audited taxpayer. At the end of the process, the IRA sends a communication to the audited taxpayer explaining the taxes and penalties due. Within this category, we further distinguish between "partial desk audits," which concern specific types of income or violations, and "full desk audits." The latter type is the most accurate, as it includes all the types of income and taxes that have been understated or omitted by the taxpayers; moreover the full desk audit can be considered final as, if nothing new emerges, IRA cannot perform any further audits.

Onsite or field audits take place at a taxpayer's premises and can be classified in two categories that we call "deep field audit" and "soft field audits"; in the case of deep field audits, the tax inspectors can stay at the taxpayer's premises for up to 30–60 working days. Tax auditors can examine all the business records or documents that are deemed relevant for the purpose of the audit. At the end of the deep field audit, the inspectors are required to produce a report with the outcome of the audit. In contrast, the soft field audits are much shorter (normally they take a few hours) and the aim is to check up on specific elements of the taxpayer's fiscal standing. A field tax audit cannot bear any request to pay taxes or fees, but all the evidence emerging from the inspection can be used by the IRA during a desk audit.

The dataset has a balanced panel structure and includes information on 126,401 taxpayers per year. Table 1 reports the number of audits over the period concerned. The total number of audited taxpayers has slightly decreased from more than 11,000 in 2006 to less than 9,000 in 2011. The analysis of the number of audits disaggregated by category shows that an increasing number of taxpayers have experienced full desk audits. Every year, on average, about 8.2 percent of the taxpayers in our sample have been audited and 1.7 percent of them were subject to a full desk audit.

TABLE 1. Number and Percentage of Italian Revenue Agency Tax Audits Among the Study Sample, by Type and Year, 2006–2011

Year	Desk Audit		Field Audit		Audited Taxpayers	Percent Audited	Total Sample
	Full	Partial	Deep	Soft			
2006	1,339	5,425	844	6,919	11,392	9.0%	126,401
2007	1,424	6,078	977	6,036	11,253	8.9%	126,401
2008	1,938	6,456	1,103	4,732	10,950	8.7%	126,401
2009	2,826	5,083	1,054	4,385	10,127	8.0%	126,401
2010	2,679	4,671	996	4,162	9,503	7.5%	126,401
2011	2,339	4,135	1,079	4,175	8,846	7.0%	126,401
Total	**12,545**	**31,848**	**6,053**	**30,409**	**62,071**	**8.2%**	**758,406**

NOTE: Some audited taxpayers have experienced more than one type of tax audit in the same year. This is why the number of audited taxpayers differs from the sum of the first four columns.

TABLE 2. Italian Revenue Agency Tax Audits by Economic Sector Among the Study Sample, 2006–2011

Economic Sector	Desk Audit		Field Audit		Audited Taxpayers	Percent Distribution		
	Full	Partial	Deep	Soft		Full Desk	All Audits	Total Sample
Sport and recreation	721	2,184	400	2,634	4,495	5.7%	7.2%	5.6%
Hotels and restaurants	1,295	4,469	873	5,030	8,142	10.3%	13.1%	5.2%
Finance and insurance	132	166	60	33	353	1.1%	0.6%	1.3%
Real estate	177	211	53	133	522	1.4%	0.8%	1.1%
Management and consulting	288	440	61	213	915	2.3%	1.5%	4.1%
Legal, accounting and consultancy services	1,535	2,241	396	1,721	5,555	12.2%	8.9%	17.6%
Wholesale trade	972	1,386	298	685	2,954	7.7%	4.8%	11.2%
Retail trade	3,124	13,656	2,252	15,782	24,579	24.9%	39.6%	19.6%
Construction	2,169	2,461	792	814	5,732	17.3%	9.2%	15.5%
Mining	1,013	2,949	582	2,690	5,424	8.1%	8.7%	7.6%
Water supply	17	21	10	9	48	0.1%	0.1%	0.1%
Education and health care	308	557	73	248	1,102	2.5%	1.8%	4.2%
Renting and travel agencies	204	335	65	170	669	1.6%	1.1%	2.0%
Information and communication technology	80	159	19	81	309	0.6%	0.5%	1.3%
Transports and storage	510	613	119	166	1,272	4.1%	2.0%	3.8%
Total	**12,545**	**31,848**	**6,053**	**30,409**	**62,071**	**100%**	**100%**	**100%**

NOTE: Some audited taxpayers have experienced more than one type of tax audit in the same year. This is why the number of audited taxpayers differs from the sum of the first four columns.

Most of the taxpayers in the sample operate in the retail and wholesale trade sectors (20 percent and 11 percent, respectively), in the legal, accounting, and consultancy sector (18 percent), and in the construction sector (15 percent) (Table 2). Interestingly, almost 40 percent of audited taxpayers operate in the retail trade sector; this is mainly due to the typical activity of inspections of shops and restaurants aimed at checking that the receipts are correctly issued and the book of accounts is correctly updated.[7] This type of inspection falls in the category of soft field audits; indeed the vast majority of soft field audits have been conducted on businesses operating in the retail trade sector and in the "hotel and restaurants" category. The distribution of full desk audits is more uniform across economic activity. However, full desk audits are relatively more frequent in the construction sector where 15.5 percent of the businesses in the sample operate and 17.3 percent of full desk audits take place.

Summary statistics for audited and nonaudited taxpayers are reported in Table 3 for 2011. The audited group includes all the taxpayers that have been audited at least once over the period of 2006–2011. Taxpayers in the nonaudited group have never been audited over the analyzed period. The trend of the taxes paid over the whole period is reported in Annex 1.

[7] See Battiston, *et al.* (2016) for more details on this type of inspection.

TABLE 3. Amounts Reported (in €) by Audited and Nonaudited Taxpayers in the Study Sample, 2011

Item	Audited	Nonaudited	Total
Number of taxpayers	38,897	87,504	126,401
Regional business tax	1,543	1,284	1,364
Value Added Tax (VAT)	1,773	1,590	1,646
Personal Income Tax (PIT)	7,910	9,329	8,892
Sales of goods and services	191,318	113,420	137,391
Labor costs	20,920	13,426	15,732
Total costs	151,327	75,177	98,610
Total revenue	210,690	122,030	149,313

All the variables (except PIT) were higher on average for the audited group, suggesting that the IRA audit selection process is not random. In particular, the IRA seems to audit larger taxpayers more often, who report higher levels of costs and revenue but, at the same time, pay less personal income tax. However, as discussed in more detail in Section 4, as long as the IRA selection process depends on some (unobservable) characteristics that are time invariant, the validity of the estimates is not affected.

Our dataset contains information on the so called *Studi di Settore* (SdS), a program to audit small businesses, introduced in Italy in 1998. Under this program, the Italian Tax Administration provides each small business (with sales below €7.5 million) with an estimated level of sales depending on the economic sector and on the reported economic inputs such as the number of employees, the cost of inputs, and so on. All the taxpayers reporting sales below the estimated level are labelled as "noncongruent." Under the SdS program, sector-specific economic indicators are also computed (such as value added per employee or change in inventories) to detect common schemes of tax evasion. Taxpayers with indicators substantially different from the computed benchmark are considered as "noncoherent."

This information is used by the IRA in the audit selection process. Indeed, failing to meet the congruence and coherence criteria implies an increased probability of receiving a tax audit.

TABLE 4. SDS Coherence and Congruence Within Study Sample, by Tax Audit Category

Item	Audited	Nonaudited	Total
Percent of SDS coherent	47%	58%	57%
Percent of SDS congruent	56%	67%	62%

Table 4 shows that 47 percent of audited taxpayers were coherent in the year when they were audited and 56 percent were congruent. These percentages are much higher in the group of nonaudited taxpayers, indicating that the failure in meeting SDS criteria dramatically increases the probability of being tax audited. Overall, 57 percent of the taxpayers meet the coherence criteria and 62 percent meet the congruence criteria. As the coherence and congruence indicators influence both the probability of tax audit assignment and the amount of taxes paid, it is very important to control for these variables in the empirical specifications to avoid endogeneity.

4. The Empirical Model

Our analysis provides evidence of the effectiveness of tax audits in promoting tax compliance. Tax enforcement activity can influence tax compliance in several ways. The existing literature has identified a direct effect and an indirect effect of tax audits on tax collection. The direct effect consists in the tax evasion directly detected, therefore resulting in additional collected revenue, sanctions, and fees. The indirect effect is mostly related to changes in the behavior of the audited taxpayer; the assumption is that taxpayers spontaneously tend

to increase the amount of income reported after having been subject to a tax inspection. Moreover, scholars have identified a third effect of auditing activity, the so-called "general deterrence effect" of audit programs. Indeed, it is plausible that when an increase in the number or in the effectiveness of tax inspection is observed, taxpayers perceive a higher risk of being audited and they therefore tend to be more careful in correctly reporting their income. This implies a reduction in the expected payoff generated by underreporting income. In the next section, we focus on the indirect effect of tax audits on tax compliance; the analysis of the general deterrence effect is described in section 4.4.

4.1. The Indirect Effect of Tax Audits

Commonly, in the policy-evaluation literature, the simplest data structure is composed of two groups (the treatment group and the control group), both of which are observed before and after the implementation of the policy. In our case, the data structure is more complicated, as we have six time periods, each firm can be audited more than once, and, in some specifications of the model, we also investigate the effect of different types of audit programs (section 4.3). Hence, we can exploit the longitudinal nature of our data by estimating the following equation with a fixed-effect model that accounts for individual time-invariant heterogeneity:

$$y_{it} = a_i + d_t + \beta A_{it} + X'_{it}\gamma + \varepsilon_{it}, \quad (1)$$

where y_{it} is the natural log of the tax paid (Regional Business Tax, VAT, or PIT depending on the specification), a_i is a time-invariant fixed effect, d_t is a taxpayer-invariant fixed effect, A_{it} is a dummy variable, which takes value 1 if taxpayer i has been audited in year t and zero otherwise, X'_{it} is a vector of control variables influencing both the probability of being audited and the tax paid, and ε_{it} is the idiosyncratic error.

Regression results are reported in Table 5 for the three taxes analyzed. On average, a generic tax inspection implies an increase in the regional business tax paid by 1.8 percent in the same year in which the inspection takes place, and by 3.1 percent in the following year. The effect of the audit is higher for the personal income tax paid, which increases by 5.2 percent in the same year and by 3.5 percent in the following year. Taxpayers seem to be able to adjust the VAT paid faster (+5.3 percent in the same year in which the audit occurs while the lagged coefficient is not statistically significant). This is probably due to the fact that the VAT can be paid monthly or quarterly; therefore a change in the behavior of audited taxpayers can be observed earlier.

TABLE 5. Tax Audit Effect on Tax Compliance

Item	Regional Business Tax	VAT	PIT
Audit	0.018	0.053	0.052
	(0.009)**	(0.016)***	(0.013)***
Audit-1	0.031	0.011	0.035
	(0.009)***	(0.016)	(0.012)***
SDS Congruence	0.304	0.281	0.433
	(0.007)***	(0.012)***	(0.009)***
SDS Coherence	0.390	0.132	0.768
	(0.007)***	(0.011)***	(0.009)***
Sales of Goods and Services	0.394	0.440	0.401
	(0.008)***	(0.007)***	(0.009)***
Total Revenue	0.140	0.017	0.211
	(0.004)***	(0.005)***	(0.006)***
Total Cost	-0.025	-0.031	-0.073
	(0.003)***	(0.004)***	(0.004)***
Labor Cost	0.038	-0.002	-0.012
	(0.001)***	(0.002)	(0.001)***
Number of Observations	633,745	650,780	644,555

** $p<0.05$; *** $p<0.01$ Time dummies included. Robust standard errors in parentheses.

4.2. Robustness of the Results

In this section we discuss whether the econometric approach used addresses the endogeneity issues that are potentially arising from the nonrandomness of the IRA selection process. Indeed, descriptive statistics suggest that the IRA audits larger businesses with higher levels of reported income and costs more frequently. The IRA selection process could also be based on some individual characteristics that we do not observe.

The model used in this paper allows us to consistently estimate the casual effect of tax audits on tax compliance, as long as the selection process of the IRA is based on time-invariant individual characteristics. Indeed, under the fixed-effect estimation procedure, a within transformation that removes the individual fixed effect is applied to the data.

However, whether there exist time-varying unobservables that are correlated both with the probability of being tax audited and with the outcome variable (*i.e.,* the tax paid, in our model) still needs to be checked. Ruling out this possibility is crucial to the validity of our results. If the IRA selection process is based on some time-varying factors that affect the tax paid, then it is impossible to determine the direction of causality. In other words the key question is the following: "Is it the tax audit that is driving up tax compliance? Or is it the case that a taxpayer was selected because of other confounding factors that also influence the tax paid?"

TABLE 6. Leads and Lags

Item	One Lead and Lag			Two Leads and Lags		
	Regional Business Tax	VAT	PIT	Regional Business Tax	VAT	PIT
Audit	0.030	0.053	0.061	0.044	0.068	0.096
	(0.011)***	(0.019)***	(0.015)***	(0.027)	(0.051)	(0.037)***
Audit $_{t-1}$	0.043	0.009	0.046	0.080	0.035	0.086
	(0.011)***	(0.019)	(0.014)***	(0.025)***	(0.048)	(0.034)**
Audit $_{t-2}$				0.035	-0.005	0.006
				(0.020)*	(0.039)	(0.027)
Audit $_{t+1}$	0.006	-0.003	0.023	0.017	-0.042	0.068
	(0.011)	(0.020)	(0.015)	(0.027)	(0.050)	(0.036)*
Audit $_{t+2}$				0.001	-0.042	0.070
				(0.022)	(0.043)	(0.031)**
SDS congruence	0.280	0.286	0.412	0.222	0.264	0.354
	(0.008)***	(0.014)***	(0.011)***	(0.014)***	(0.026)***	(0.019)***
SDS coherence	0.447	0.118	0.775	0.529	0.178	0.918
	(0.008)***	(0.013)***	(0.010)***	(0.013)***	(0.025)***	(0.018)***
Sales of goods and services	0.442	0.463	0.443	0.428	0.518	0.445
	(0.010)***	(0.009)***	(0.011)***	(0.014)***	(0.016)***	(0.017)***
Total revenue	0.144	0.022	0.207	0.144	0.022	0.218
	(0.005)***	(0.006)***	(0.007)***	(0.008)***	(0.011)**	(0.011)***
Total costs	-0.032	-0.032	-0.068	-0.038	-0.039	-0.072
	(0.003)***	(0.005)***	(0.004)***	(0.006)***	(0.009)***	(0.007)***
Labour costs	0.035	-0.003	-0.009	0.028	-0.006	-0.004
	(0.002)***	(0.002)	(0.002)***	(0.002)***	(0.004)	(0.003)
Number of observations	505,604	520,624	515,644	252,802	260,312	257,822

* $p<0.1$; ** $p<0.05$; *** $p<0.01$ Time dummies included. Robust standard errors in parentheses.

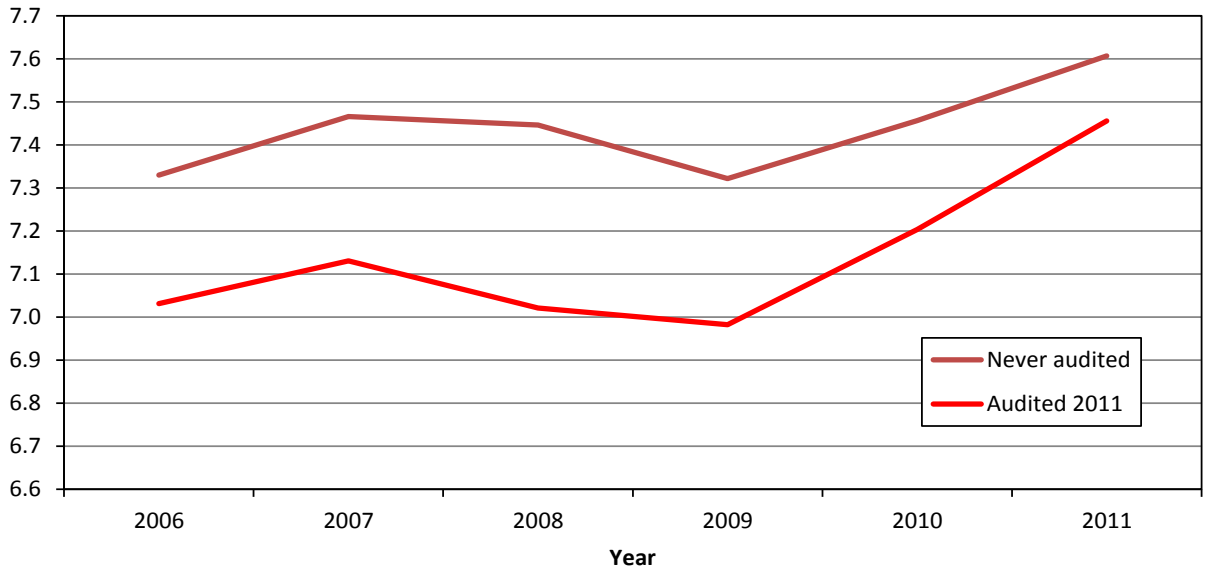
Results of the regression are reported in Table 6. Estimating the model with one lead and one lag, we find that future audits do not have any statistically significant explanatory power on the tax paid. However, running the same regression using two leads and two lags of the audit variable, the consistency of the estimates is confirmed for the Regional Business Tax and for the VAT model, but not for the PIT model.

To answer this question we use an approach originally developed by Autor (2003). We estimate equation (1) by introducing leads and lags of the audit variable.[8] We expect that past audits help to predict tax paid, while future audits do not.[9] When audits have an anticipatory effect on income reported and tax paid, it is not possible to determine the direction of causality (this is what Angrist and Pischke (2008) call a Granger test, used "to see whether causes happen before consequences and not vice versa").

Therefore, we further investigate the validity of the estimates of the PIT model by analyzing the trend of the PIT declared. Indeed, a crucial assumption for the validity of the estimates is the so-called common trend assumption. The hypothesis is that, without the treatment, the control group and the treatment group would have shared the same trend of the outcome variable. More specifically, in our case, the condition for the validity of the estimates requires that the pre-audit trend of the audited taxpayers is the same as the trend of the nonaudited taxpayers group.

Figure 1 compares the value of the average PIT declared by year (in logs) in the group of taxpayers audited in 2011, with the average PIT declared by taxpayers who were never audited in the analyzed period.[10]

FIGURE 1. Parallel Trend—Log of Personal Income Tax Declared



The trend is very similar across the two groups in the pretreatment period. The average PIT declared by the group of taxpayers who were audited in 2011 is always below the average PIT declared by the nonaudited taxpayers. However, as expected, in 2011 the amount of PIT declared by the audited taxpayers increases, reducing the difference between the groups.

The same analysis, considering (shorter) pretreatment periods for the group of taxpayers audited in previous years, leads to a similar result.

4.3. Tax Audit Effect by Type of Tax Audit

As the IRA carries on different audit programs, disentangling the causal channels through which audits exert their effect can be extremely relevant when forming policy advice. Depending on where the inspections actually take place, we identify two broad categories of tax audits—desk audits and field audits. These categories

[8] See Angrist and Pischke (2008) for a detailed discussion.

[9] The same approach to test for strictly endogeneity is proposed by Wooldridge (2002).

[10] In this graphical analysis, we drop all the taxpayers who had been audited before 2011.

can be further broken down, depending on the accuracy of the audit and on the type of tax analyzed, as described in section 3. We end up with four groups of audits: (1) full desk audits; (2) partial desk audits; (3) deep field audits; and (4) soft field audits. In the same year, taxpayers could be subject to different kinds of audits on different types of income.

We estimate equation 1, including in the set of explanatory variables the four categories of tax audits described above (Table 7).

TABLE 7. Tax Audit Effect on Tax Compliance, by Type of Audit

Item	Regional Business Tax (1)	Regional Business Tax (2)	VAT	PIT
Full desk audit $_{t}$	0.190	0.189	0.181	0.147
	(0.019)***	(0.019)***	(0.032)***	(0.026)***
Partial desk audits $_{t}$	-0.009	-0.007	-0.019	0.013
	(0.014)	(0.014)	(0.024)	(0.018)
Field audit deep $_{t}$	0.008	0.008	0.022	0.037
	(0.025)	(0.025)	(0.047)	(0.039)
Field audit soft $_{t}$	-0.032	-0.079	0.046	0.013
	(0.014)**	(0.022)***	(0.025)*	(0.020)
Full desk audit $_{t-1}$	0.176	0.175	0.109	0.119
	(0.019)***	(0.019)***	(0.034)***	(0.027)***
Partial desk audits $_{t-1}$	0.023	0.024	0.011	0.029
	(0.013)*	(0.013)*	(0.023)	(0.018)*
Field audit deep $_{t-1}$	0.047	0.047	-0.020	-0.026
	(0.025)*	(0.025)*	(0.047)	(0.040)
Field audit soft $_{t-1}$	-0.048	-0.098	-0.029	0.017
	(0.013)***	(0.019)***	(0.024)	(0.018)
SDS congruence	0.302	0.299	0.280	0.433
	(0.007)***	(0.007)***	(0.012)***	(0.009)***
SDS coherence	0.389	0.388	0.132	0.768
	(0.007)***	(0.007)***	(0.011)***	(0.009)***
Sales of goods and services	0.407	0.407	0.440	0.401
	(0.008)***	(0.008)***	(0.007)***	(0.009)***
Total revenue	0.141	0.141	0.017	0.211
	(0.004)***	(0.004)***	(0.005)***	(0.006)***
Total costs	-0.025	-0.025	-0.031	-0.073
	(0.003)***	(0.003)***	(0.004)***	(0.004)***
Labor costs	0.038	0.038	-0.002	-0.012
	(0.001)***	(0.001)***	(0.002)	(0.001)***
Field audit soft and SDS congruence $_{t}$		0.081		
		(0.025)***		
Field audit soft and SDS congruence $_{t-1}$		0.091		
		(0.022)***		
Number of observations	632,005	632,005	650,780	644,555

* $p<0.1$; ** $p<0.05$; *** $p<0.01$ Time dummies included. Robust standard errors in parentheses.

As expected, the results show that full desk audits are the most effective in increasing tax compliance. In particular, this type of inspection increases the regional business tax declared by about 19 percent in the same year when the audit takes place and by 17 percent in the following year, the VAT declared by about 18 percent and 11 percent, respectively, and the PIT declared by 14.7 percent and 12 percent. The direct effect of the other types of tax audits on tax compliance is more limited.

Surprisingly, the results reported in the first column of Table 7 show a statistically significant negative effect of the soft field audits on tax compliance. However, as taxpayers who are "noncongruent" (taxpayers reporting a level of income substantially below the level estimated by the SdS) are much more often subject to soft field audits, we investigate whether the effect of this type of audit is different across congruent and noncongruent taxpayers. Results reported in column 2 show that the effect of soft field audits is positive on the sample of congruent taxpayers. Indeed, taxpayers know that reporting income below the level estimated by SdS dramatically increases their probability of being audited. Therefore, it might be the case that noncongruent taxpayers are on average already compliant and do not adjust their level of declared income as a consequence of a tax inspection.

Even though field audits exert a statistically weaker effect on tax compliance, they are a very important part of the overall IRA strategy to increase tax compliance. Indeed, they often serve as input to subsequent desk audits.

4.4. The General Deterrence Effect

The question of whether a general deterrence effect exists has been discussed widely in the literature, as the empirical evidence has not been conclusive on the relationship between audit rates and tax compliance. In principle, increasing the number of audits (measured by audit rates) implies an increase in the probability of being audited, and should thereby incentivize tax compliance. Likewise, the effectiveness of the tax authorities in detecting tax evasion should exert a significant effect on tax compliance. Both the probability of being tax audited and, once audited, the ability of the inspectors should reduce the expected gain of tax evasion. While the probability of being audited could be easily approximated using the audit rate, it is more difficult to find a good proxy for the average effectiveness of the IRA in detecting evasion.

To measure if the IRA activity exerts a deterrence effect that spills over to nonaudited taxpayers, we include in equation 1 the audit rate by province and the ratio of the additional regional business tax evaded and detected by the IRA to the overall regional business tax declared by province. The latter variable measures both the probability of being audited and the amount of tax evasion detected on the total tax declared.[11]

As shown in Table 8, if the audit rate in the province where the taxpayer is resident for tax purposes increases by 1 percent, tax compliance increases between 1.1 percent and 1.7 percent. The effect of the regional business tax evaded and detected to the overall amount of tax declared is also positive and statistically significant as expected.

[11] There can be some simultaneity; more tax compliance in one province can induce tax administration to decrease the number of audits in that province while, at the same time, more tax audits in that province can have a positive effect on tax compliance. The econometric approach used in this paper allows us to address the endogeneity caused by the correlation between the error term and any time invariant individual unobservable. So, even if the revenue agency sets the audit rates in each province using a criterion based on some time invariant characteristic (which is correlated with the compliance), then our estimates are still consistent and unbiased. Moreover, both audit rates and the additional regional business tax evaded and detected by the IRA, as well as the overall regional business tax declared by province depend on the aggregate tax compliance at the province level but our dependent variable is disaggregated at the individual level. This could partly mitigate the endogeneity problem.

TABLE 8. General Deterrence (Spillover) Effect, Regional Business Tax

Item	Model 1	Model 2	Model 3
Full desk audit $_{t}$	0.135	0.138	0.135
	(0.023)***	(0.023)***	(0.023)***
Partial desk audits $_{t}$	-0.012	-0.011	-0.011
	(0.018)	(0.018)	(0.018)
Field audit deep $_{t}$	-0.007	-0.007	-0.007
	(0.032)	(0.032)	(0.032)
Field audit soft $_{t}$	-0.022	-0.022	-0.022
	(0.018)	(0.018)	(0.018)
Full desk audit $_{t-1}$	0.130	0.129	0.129
	(0.023)***	(0.023)***	(0.023)***
Partial desk audits $_{t-1}$	-0.005	-0.004	-0.004
	(0.016)	(0.016)	(0.016)
Field audit deep $_{t-1}$	0.026	0.026	0.026
	(0.031)	(0.031)	(0.031)
Field audit soft $_{t-1}$	-0.013	-0.012	-0.012
	(0.016)	(0.016)	(0.016)
Percent additional regional business tax evaded and detected by province		0.033	0.032
		(0.004)***	(0.004)***
Audit rate by province	0.017		0.011
	(0.005)***		(0.005)**
SDS congruence	0.275	0.274	0.274
	(0.009)***	(0.009)***	(0.009)***
SDS coherence	0.479	0.479	0.479
	(0.009)***	(0.009)***	(0.009)***
Sales of goods and services	0.362	0.362	0.362
	(0.009)***	(0.009)***	(0.009)***
Total revenue	0.143	0.143	0.143
	(0.006)***	(0.006)***	(0.006)***
Total costs	-0.029	-0.029	-0.029
	(0.004)***	(0.004)***	(0.004)***
Labor costs	0.033	0.033	0.033
	(0.002)***	(0.002)***	(0.002)***
Number of observations	419,072	419,072	419,072

** $p<0.05$; *** $p<0.01$ Time dummies included. Robust standard errors in parentheses.

5. Conclusions

Building and innovating on a companion paper (D'Agosto, et al. 2017), this study extends the analysis of the effect of tax audits on tax compliance across different types of taxes and audit programs. Along the lines of previous research, our results confirm the evidence that tax inspections increase taxpayers compliance.

Findings suggest that the relative magnitude of the effect is different across taxes; an audit increases the regional business tax declared by 1.8 percent, the VAT by 5.3 percent, and the PIT by 5.2 percent. Likewise, the persistence of the effect of the audit over time also differs across type of tax. While for both regional business tax and personal income tax the effect remains positive also in the first year after the audit, in the case of VAT this effect seems to vanish in the first year following the audit.

Full desk audits turn out to be the most effective in increasing tax compliance. In particular, taxpayers who experienced a full desk audit increased the regional business tax they declared by about 19 percent in the year of the audit, the VAT declared by about 18 percent, and the PIT by 14.7 percent. Field audits seem to exert a weaker effect on tax compliance. However, as they often serve as input to subsequent desk audits, their effect may be underestimated. It remains for future research to better understand the interactions between different types of audits.

Finally, we present a preliminary analysis of the general deterrence effect of tax audits on tax compliance. We proxy the probability of being audited and the effectiveness of tax inspectors in detecting tax evasion with the audit rate and with the ratio of the tax evaded and detected to the overall amount of tax paid, both disaggregated by province. Both variables exert a statistically significant effect on tax compliance.

Annex 1. Dynamic of Average Taxes Declared for the Audited and Nonaudited Groups

FIGURE A1. Number of Audited and Nonaudited Taxpayers—Regional Business Tax

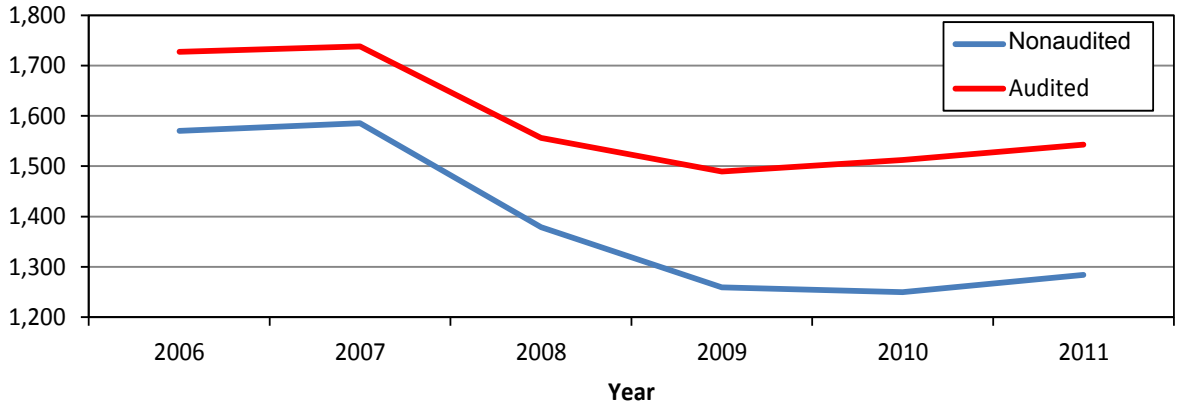


FIGURE A2. Number of Audited and Nonaudited Taxpayers—VAT

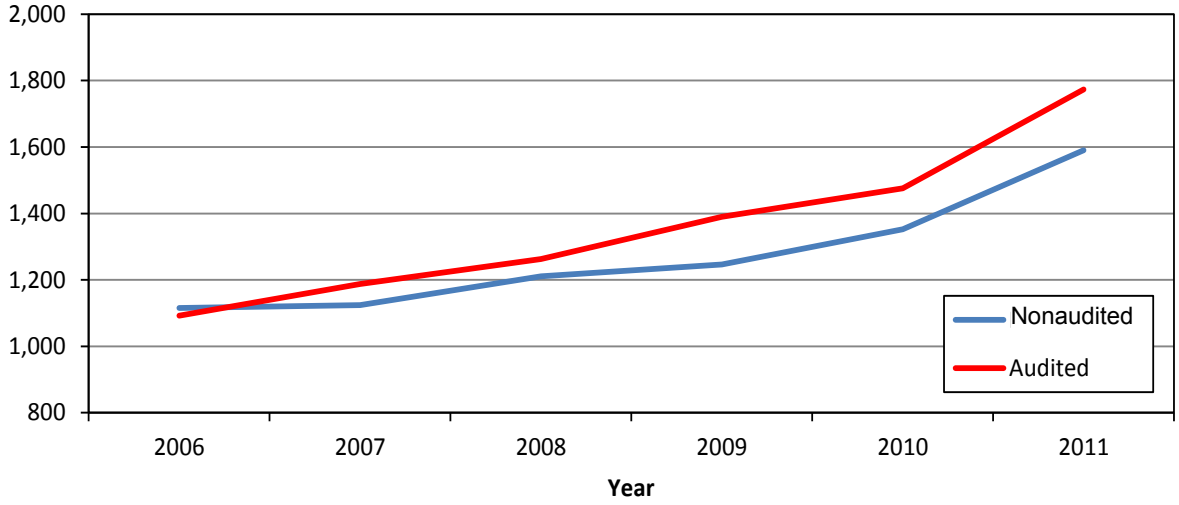
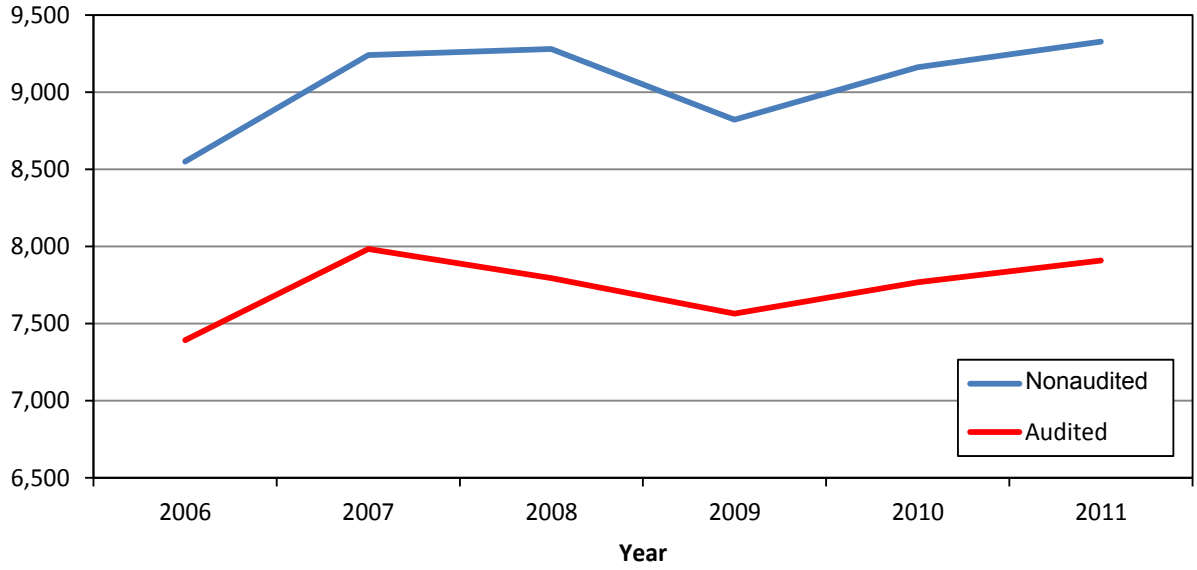


FIGURE A3. Number of Audited and Nonaudited Taxpayers—PIT



References

Allingham, Michael G., and Agnar Sandmo. 1972. Income Tax Evasion: A Theoretical Analysis, *Journal of Public Economics*. 1: 323–338.

Alm, James. 1988. Compliance Costs and the Tax Avoidance-Tax Evasion Decision. *Public Finance Quarterly.* 16: 31–66.

Alm, James, and Michael McKee. 2004. Tax Compliance as a Coordination Game. *Journal of Economic Behavior & Organization* 54(3): 297–312.

Alm, James, Betty R. Jackson, and Michael McKee. 2009. Getting the Word Out: Enforcement Information Dissemination and Compliance Behaviour. *Journal of Public Economics* 93(3–4): 392–402.

Alm, James, Erich Kirchler, and Stephan Muehlbacher. 2012. Combining Psychology and Economics in the Analysis of Compliance: From Enforcement to Cooperation. *Economic Analysis and Policy* 42(2): 133–152.

Andreoni, James, Brian Erard, and Jonathan Feinstein. 1998. Tax Compliance. *Journal of Economic Literature* 36(2): 818–860.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Autor, David H. 2003. Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics* 21(1): 1–42.

Battiston, P., D. Duncan, S. Gamba, and A. Santoro. 2016. The Italian Blitz: A Natural Experiment on Audit Publicity and Tax Compliance. FBK-IRVAPP Working Paper.

Becker, Gary S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76: 169–217.

Birskyte, Liucija. 2013. Effects of Tax Auditing: Does the Deterrent Deter? *Research Journal of Economics, Business and ICT* 8(2).

Blumenthal, Marsha, Charles Christian, Joel Slemrod, and Matthew G. Smith. 2011. Do Normative Appeals Affect Tax Compliance? Evidence From a Controlled Experiment in Minnesota. *National Tax Journal* 54(1): 125–38.

D'Agosto, Elena, Massimiliano Marigliani, and Stefano Pisani. 2013. Asymmetries in the Territorial VAT Gap. Italian Revenue Agency, Discussion Topics, n.2.

D'Agosto, Elena, Marco Manzo, Stefano Pisani, and Filippo M. D'Arcangelo. 2017. The Effect of Audit Activity on Tax Declaration: Evidence on Small Businesses in Italy, *Public Finance Review*.

Davis, Jon S., Gary Hecht, and Jon D. Perkins. 2003. Social Behaviors, Enforcement and Tax Compliance Dynamics. *The Accounting Review* 78(1): 39–69.

Di Porto, Edoardo. (2011). Undeclared Work, Employer Tax Compliance, and Audits. *Public Finance Review* 39(1): 75–102.

Dubin, Jeffrey A., and Louis L. Wilde. 1988. An Empirical Analysis of Federal Income Tax Auditing and Compliance. *National Tax Journal* 41: 61–74.

Dubin, Jeffrey A., Michael J. Graetz, and Louis L. Wilde. 1990. The Effect of Audit Rates on the Federal Individual Income Tax, 1977–1986. *National Tax Journal* 43: 395–409.

Dubin, Jeffrey A. 2007. Criminal Investigation Enforcement Activities and Taxpayers Noncompliance. *Public Finance Review* 35(4): 500–529.

Erard, Brian. 1992. The Influence of Tax Audits on Reporting Behavior. In *Why People Pay Taxes: Tax Compliance and Enforcement*, Joel Slemrod, 95–114, The University of Michigan Press.

Fiorio, Carlo V., Stefano Maria Iacus, and Alessandro Santoro. 2013. Taxpaying Response of Small Firms to an Increased Probability of Audit: Some Evidence from Italy. University of Milan Bicocca Department of Economics, Management and Statistics Working Paper 251.

Gemmell, Norman, and Marisa Ratto. 2012. Behavioral Responses to Taxpayers Audits: Evidence from Random Taxpayer Inquiries. *National Tax Journal* 65(1): 33–58.

Harmizade, Nigar, Gareth D. Myles, and Binh Tran-Nam. 2012. Application of Behavioural Economics to Tax Evasion. *Journal of Economic Surveys* 27(5): 941–977.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez. 2010. Unwilling or Unable To Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, Volume 79, Issue 3.

Manestra, Stefano. 2010. A Short History of Tax Compliance in Italy. Bank of Italy Occasional Papers 81.

Niu, Yonghzhi. 2011. Tax Audit Impact on Voluntary Compliance. *Journal of Economic and Social Measurement* 36(4): 237–251.

OECD. 2009. *Managing and Improving Compliance: Recent Developments in Compliance Risk Treatments*. Forum on Tax Administration: Compliance Sub-Group, Information Note, Centre for Tax Policy and Administration.

Plumley, Alan Howard. 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).

Plumley, Alan Howard. 2002. The Impact of the IRS on Voluntary Tax Compliance: Preliminary Empirical Results. National Tax Association, 95th Annual Conference on Taxation Orlando, Florida, November 14–16, 2002.

Ratto, Marisa, Richard Thomas, and David Ulph. 2013. The Indirect Effects of Auditing Taxpayers. *Public Finance Review* 41(3): 317–333.

Reinganum, Jennifer, and Louis L. Wilde. 1986. Credibility and Law Enforcement. California Institute of Technology, Division of the Humanities and Social Sciences, Working Paper 604.

Rincke, Johannes, and Christian Traxler. 2011. Enforcement Spillovers. *Review of Economics and Statistics* 93(4):1224–1234.

Santoro, Alessandro. 2008. Taxpayers' Choices Under Studi Di Settore: What Do We Know And How We Can Interpret It?". *Giornale degli Economisti e Annali di Economia* 67(2): 161–184.

Santoro, Alessandro, and Carlo V. Fiorio. 2011. Taxpayer Behavior When Audit Rules Are Known: Evidence from Italy. *Public Finance Review* 39(1): 103–123.

Santoro, Alessandro. 2016. Do Small Businesses Respond to a Change in Tax Audit Rules? Evidence from Italy. *Public Finance Review* (2016): 1091142116676361.

Scotchmer, Suzanne, and Joel Slemrod. 1989. Randomness in Tax Enforcement. *Journal of Public Economics* 38(1): 17–32.

Slemrod, Joel. 1991. *Why People Pay Taxes: Tax Compliance and Enforcement*. The University of Michigan press.

Slemrod, Joel. 2007. Cheating Ourselves: The Economics of Tax Evasion. *The Journal of Economic Perspectives* 21.1 (2007): 25–48.

Tagkalakis, Athanasios O. 2014. The Direct and Indirect Effects of Audits on the Tax Revenue in Greece. *Economics Bulletin* 34(2): 984–1001.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Valuing Unpaid Tax Assessments: Estimating Long-Run Collectability Using an Econometric Approach

Alex Turk, with Eric Henry, Dan Howar, and Maryamm Muzikir (Internal Revenue Service)[1]

Introduction

The Government Management Reform Act of 1994[2] and the Chief Financial Officers (CFO) Act of 1990[3] made the Government Accountability Office (GAO) responsible for annual audits of agency-wide financial statements and the U.S. government's consolidated financial statements. As of Fiscal Year 2016, there were 400 billion dollars in unpaid assessments (UA) the IRS has the authority to collect.

There are four financial classifications of UA inventory assets. For this paper, just like in the financial statements, the memo financial classification will be excluded, as the assessments in this inventory are not true receivables. It does not meet any of the other financial classifications as it contains fraudulent/frivolous tax returns and multi-year examinations with global netting issues. The remaining classifications are:

- **Taxes receivable**—assessments that the taxpayer has agreed to, or there has been a legal determination of liability;
- **Compliance assessments**—unagreed enforcement assessments; and
- **Write-offs**—amounts deemed to have little collection potential, but by statute must remain on the books for the length of the collection statute (usually 10 years).

The IRS reports on its financial statements an estimated dollar amount it expects to collect from the taxes receivable portion of the unpaid assessments inventory. Obtaining a clean GAO audit rests in part on the accuracy of the UA collectability estimate. In this paper, we develop an econometric approach to estimating the value of unpaid assessments inventory. This approach is data-driven and provides a more comprehensive estimate of the value of the entire inventory of unpaid assessments over the life of each asset.

Unpaid Assessments and Collectability Estimates

While a major component of the IRS's mission is assessing and collecting the proper amount of tax, the challenge for IRS financial management is to accurately account for and determine the net realizable value of an ever-increasing inventory of unpaid assessments. This can be challenging in an environment with increasingly limited resources. The IRS has long sought a more flexible decision tool that could strengthen its financial reports and reconcile reports to downstream collection activities driven by an individual debtor's characteristics.

Unpaid assessments consist of taxes, penalties, and interest that have not been collected or abated. IRS CFO staff are responsible for estimating, reconciling, analyzing trends, and preparing projections on the unpaid assessment inventory. These tasks require personnel to evaluate taxpayer account information over time to identify those characteristics that impact the probability of collections. Reconciliation of balances to unpaid assessment information reports ensures the integrity of the financial statements.

---

[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] Civic Impulse. (2017). S. 2170—103rd Congress: Government Management Reform Act of 1994. Retrieved from https://www.govtrack.us/congress/bills/103/s2170.

[3] Chief Financial Officers (CFO) Act of 1990 (1990, Nov). Public Law 101-576, 101st Congress. Retrieved from http://www.gao.gov/special.pubs/af12194.pdf.

Federal accounting standards (FASAB 7) require reporting Federal taxes receivable net of allowance for loss on taxes receivable on the balance sheet and notes to the financial statements, and the disclosure of compliance assessments and write-offs in the supplemental information to financial statements.[4] New requirements for the accounting of expected credit losses are set forth in IFRS 9 (International Financial Reporting Standard) and take effect for large banks' 2018 financial statements.[5] Accounting principles in the U.S. have evolved over the last 80 years, but they are still considered to be more rules-based in their approach to accounting standards and may not address unforeseen issues that arise in the normal course of business. Some government entities including the U.S. Postal Service, U.S. Department of Treasury entities, and smaller executive and legislative branch entities continue to apply this rules-based FASB (Financial Accounting Standards Board) approach. Principles-based standards, such as Generally Accepted Accounting Principles (GAAP) from the FASAB, provide more flexible applications to a broad range of situations.[6] More than 100 countries have adopted IFRS to some degree. FASAB now requires a loss allowance based on the expected losses over the life of the assets.

The recovery model-based approach presented in this paper supports the process of estimating recovery of taxes receivable with the corresponding allowance for loss (loss given default). Scores developed from the models provide an opportunity to automate and enhance the annual financial statement audit by reducing the inherent risk associated with small sample designs, thereby helping to ensure the IRS receives an unqualified (clean) opinion.

For most unpaid assessments, the statutory period for collection is 10 years. During the statutory period, changes in economic conditions, tax law, tax administration policy, and resources devoted to tax administration can potentially affect the actual collection from some assets. The current method for estimating the value of unpaid assessments involves conducting intensive reviews of a sample of the inventory to determine collectability. The IRS currently uses subject-matter experts to confirm that the assessments are classified properly in the sample, calculates the value of the taxes receivable portion of the sample, and then projects potential collection to the taxes receivable population of the entire UA inventory. This figure is reported in the IRS financial statements as the net realizable value of the UA inventory.

Under the current sampling method, the accuracy of the reported UA inventory amount relies heavily on the subject-matter expert's classification of the assessments in the sample. This process is very labor intensive. The fact that the process relies on the financial classification poses an inherent risk. If programming changes are made in the IRS business systems that are not reflected/recognized in the financial systems, UA assets could be misclassified. When assessments are misclassified in the audit sample, that error is projected to the population estimate. This can put the IRS at risk of not obtaining a clean audit opinion on its financial statements from GAO. In addition, the method does not value assets where the taxpayer has not agreed with the IRS's assessment or those that have been moved into a "Write-off" financial classification. Furthermore, this method does not account for policy and economic changes potentially affecting collectability over the life of the asset.

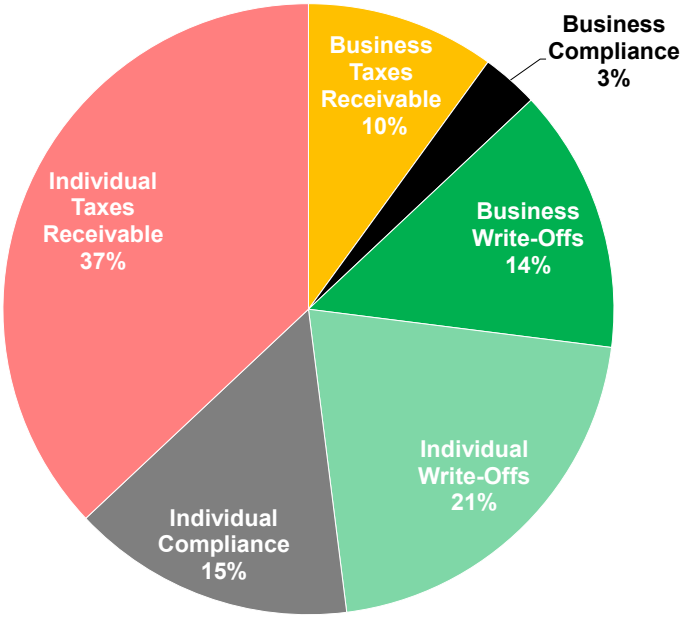
Figure 1 shows the composition of the FY 2016 unpaid assessments by financial classification. Taxes receivable assets are 47 percent of the UA inventory. The current sample method for valuing UA does not account for the potential collectability of the remaining 53 percent of unpaid assessments that are classified as either write-offs or compliance assessments.

[4] Federal Accounting Standards Advisory Board (FASAB).

[5] GPPC (2016)

[6] http://www.fasab.gov/appropriate-source-of-gaap/.

FIGURE 1. Composition of the Total $400 Billion of Gross Unpaid Assessments



SOURCE: Custodial Detail Database (CDDB) as of September 30, 2016, Individual Master File and Business Master File extract cycle 201637.

Objective

In this research, we develop a predictive model to determine the net realizable value of each unpaid assessment at any point in time. The model enables us to estimate the proportion of the current balance that will be recovered over its remaining life. We define "recovery" as the total net payments to be realized in the future as a percentage of the current module balance. The model estimates the amounts to be collected against the current balance of each UA asset based on an estimated proportion that will be recovered.

Research Design: Recovery Model

Overview

We use a logistic model to estimate the recovery rate—the proportion of the current unpaid balance that will be paid over the remaining life of the statute. The data are compiled from the IRS Compliance Data Warehouse (CDW) Unpaid Assessments data, referred to as ARDI (Accounts Receivable Dollar Inventory) in the CDW. We pool repeated January cross sections of the inventory from 2004 to 2014. We compile annual net payments from 2002 to 2016. The models provide scores or estimates of the percentage of the current balance the IRS can expect to recover on each asset in the Individual and Business Master File unpaid assessment inventory. A two-step modeling approach is used to statistically control for cases where all potential payments are not observed. These controls allow us to back out the impact of only partially observing payment streams on some assets, thus allowing a "full statute" payment/recovery estimate. Eight model specifications are created, based on the Master File Tax Class of the unpaid assessments assets.

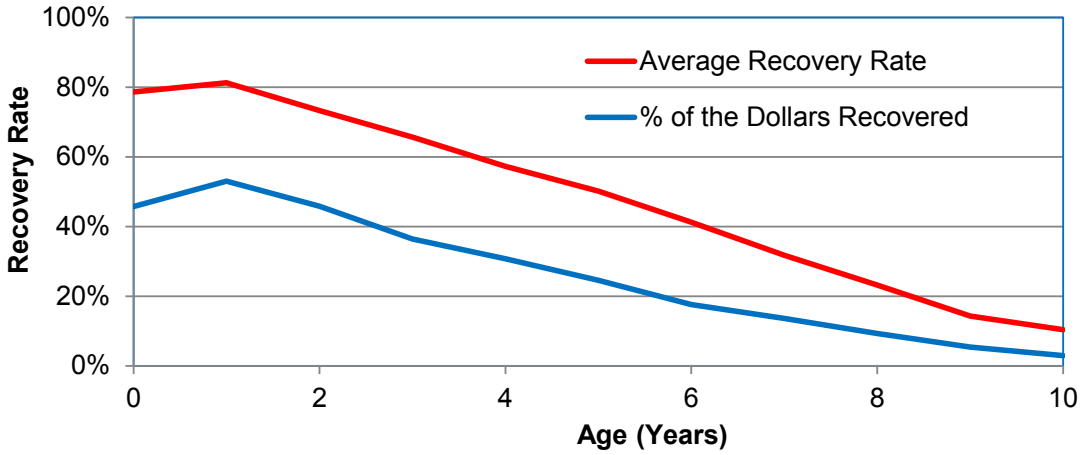
Data Dilemma: Incomplete Payments vs. Data Currency

Most UA assets have 10-year statutes, and at any point in time the UA inventory contains assessments from multiple years. Unless you go back in time more than 10 years, there will be a proportion of cases where all the potential payments are not observed. This creates a data dilemma. Typically, the most recent data are more relevant for prediction because they reflect today's economic, resource, and policy situations. However, looking at inventories further back in time allows for more cases where all the payments have been observed.

Also, having multiple years of UA inventory allows for more variation in the mix of cases and business cycle fluctuations.

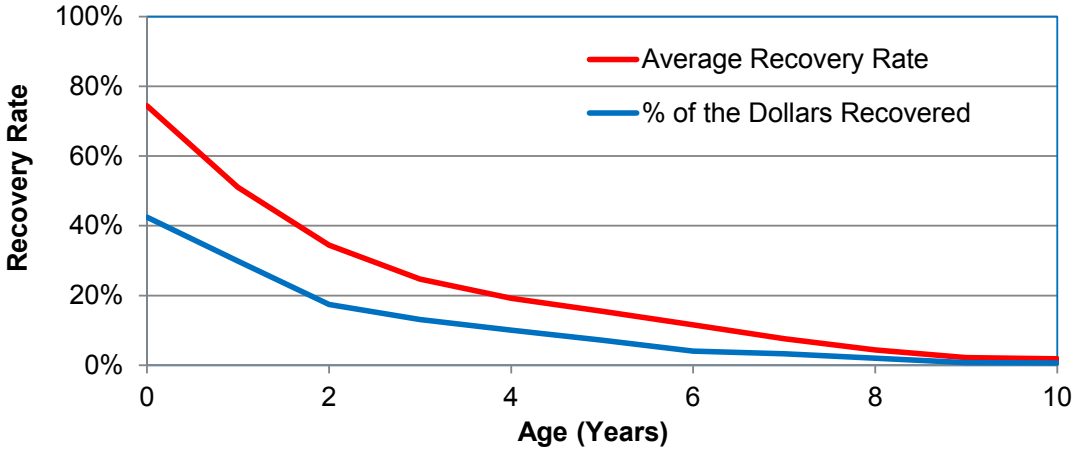
Figures 2 and 3 show recovery rates by the age of the assessment for Form 1040 and Form 941 (respectively) for the January 2015 UA inventory. In general, the recovery rates for both forms are lower for older assessments. The decay in recovery rate is much slower for individuals than for businesses. These trends reflect the fact that many assessments are resolved early in the collection process. Assessments that are 10 years old make up a smaller proportion of UA collections compared to younger cases for both individual and business tax classes. Over time, debt becomes less collectable, and less collectable debt tends to get older. Both facts contribute to the lower observed recovery for older debt.

FIGURE 2. Recovery Rate for the 2005 Inventory of Individual Master File Form 1040 Unpaid Assessments



SOURCE: Compliance Data Warehouse, Unpaid Assessment Entity and Module information, Unpaid Assessment inventory as of January 2005, net payments on the associated modules 2005–2015.

FIGURE 3. Recovery Rate for the 2005 Inventory of Business Master File Form 941 Unpaid Assessments



SOURCE: Compliance Data Warehouse, Unpaid Assessment Entity and Module Information, Unpaid Assessment inventory as of January 2005, net payments on the associated modules 2005–2015.

The recovery model needs to account for the truncated payments observed for the most recent UA assets to properly model the relationship between time and asset collectability. This truncation is more difficult to account for because the data are a repeated cross section of UA inventory. As such, separate logistic models were used to approximate the likelihood that all the potential payments are not observed by the end of the sample period. The predicted probability is included in the recovery model as a control for payment truncation on recovery estimates. In addition, we include a control for time remaining on the collection statute at the end of our sample.

Controlling for Unobserved Payments

An asset has unobserved payments when at the end of our sample period the case is still in UA and the observed recovery is less than 100 percent. Future payments can and do exceed the current balance in many cases since interest and penalties will continue to accrue. The first-stage model is designed to calculate the probability of unobserved payments on an asset.

Let T =1 if the above conditions are met and zero otherwise. Then model T as:

$$\textbf{\textit{Prob(T=1)}} = \frac{e^{Z\alpha}}{1+e^{Z\alpha}},$$

where Z contains variables in *X* (defined below) and year dummy variables. These annual dummies control for the timing of the observations and thus how long payments can be observed. This functional form provides probabilities for the payment stream being truncated. We include this probability in the recovery model as an additional explanatory variable to control for assets having unobserved payments.

Recovery Model

Consider the following basic recovery model form:

Let P_t be the net payments made during a year on an unpaid assessment module and B_n be the current total module balance, then define the variable *Y* as:

$$Y = \frac{\Sigma_{t=n}^{10} P_t}{B_n}.$$

Then let $R = \text{Max}(Y,1)$.

Then, the estimated recovery model becomes:

$X_t\beta = \beta_0 + \beta_1 ln\ (Module\ Balance) + \beta_2\ ln\ (Entity\ Balance) + \beta_3 Age + \beta_4 Age^2 + \beta_5 (Net\ Payment\ in\ Prior\ Year/ Module\ Balance) + \ldots$

$+ \beta_p\ (Probability\ of\ not\ observing\ all\ payments)$

$+ \beta_T\ (Time\ Remaining\ on\ the\ statute\ at\ the\ end\ of\ the\ sample).$

We can then model recovery, *R*, as

$$Estimated\ Recovery = R = F(X\beta) = \frac{e^{X\beta}}{1+e^{X\beta}},$$

where *F( )* is a cumulative logistic distribution function. The nonlinear logistic regression model will provide the expected recovery, ranging between zero and one, on each asset given the asset's array of characteristics, *X*.

For a "full statute" estimate of recovery, we set:

- The probability of not observing all payments = 0 and,
- Time remaining on the statute at the end of the sample = 0.

As previously mentioned, separate models and calculations are estimated for individual and business tax classes. The dependent variable "recovery" is defined as total net payments in the future as a percentage of the current balance. Recovery is considered as 100 percent when the accumulated payments exceed the current balance. Payments are not discounted based on when they are received.[7] The model generates an estimated recovery percentage that ranges from 0 to 100 percent.

The models control for measures of the taxpayer's income, number of unpaid assessment modules the taxpayer has in UA inventory, the age of the assessments, payments in each of the prior 2 years, and current location and status in the collection process. The models also control for major source of assessments and transaction category, financial classification, previous filing compliance, and type of taxpayer.

Estimated value is then calculated as the expected percent recovered multiplied by the current balance. This estimates the dollar amount the IRS expects to collect from the balance today over the remaining life of the asset. Unlike the sample method that produces estimates for only taxes receivable assets in UA inventory, the recovery models produce estimates for all assets regardless of financial classification.

Results

The results for separate model specifications were estimated for the various Individual and Business tax classes/form types. Table 1 shows the parameter estimates for the variables used to control for truncated payments in the models: "probability of truncated payments" and "time remaining at the end of the sample." In all but one instance, the "probability of truncated payments" estimates are positive and "time remaining at the end of the sample" estimates are negative. Thus, these controls work in opposite directions when the variables are set to zero in the full statute estimate calculation. This at first would seem counter-intuitive. However, the estimates are the result of pooling multiple cross sections of data over multiple years, each with different periods of time to observe payments. The more time remaining to observe payments at the end of the sample, the larger the adjustment that occurs to the estimated recovery rate. However, that adjustment is counterbalanced by the likelihood that the payments are truncated. This probability is related to the collectability of the case. So, for a given cross section of data, observations that are less collectible get a smaller "bump" in estimated recovery. That bump will vary in the sample depending on the yearly cross section from which the case originates.

Estimates from GAO audit samples show collections have nearly doubled ($26.3 billion to $49.2 billion) from 2008 to 2016, far outpacing the rate of growth in adjusted taxes receivable ($124.3 billion to $178.4 billion).[8] Figure 4 shows the model predicted average and actual observed recovery rates and recovered dollars for the Individual Income Tax Form 1040 UA inventory in Calendar Years January 2004 through 2016. In Figure 4, actual observed dollars collected on individual income tax UA are shown at nearly 30 percent, declining for more recent years because in these years, collections on more recent assets have not yet been realized. The difference between the estimated dollar recovery rate and the observed dollar recovery rate illustrates the effect of payment truncation. The more constant rate shown in the estimated dollar recovery line is the result of estimating the yet-to-be observed collections on the newer assets. More information on the predictions is provided in the appendix. This effect is shown by the increasing percent-truncated line over time. As more cases have truncated payment streams, the actual and estimated get further apart. If we don't back out the impact of the truncation, the predicted rates tend to follow the observed.

[7] The current UA valuation process does not make any attempt to account for the timing of payments.

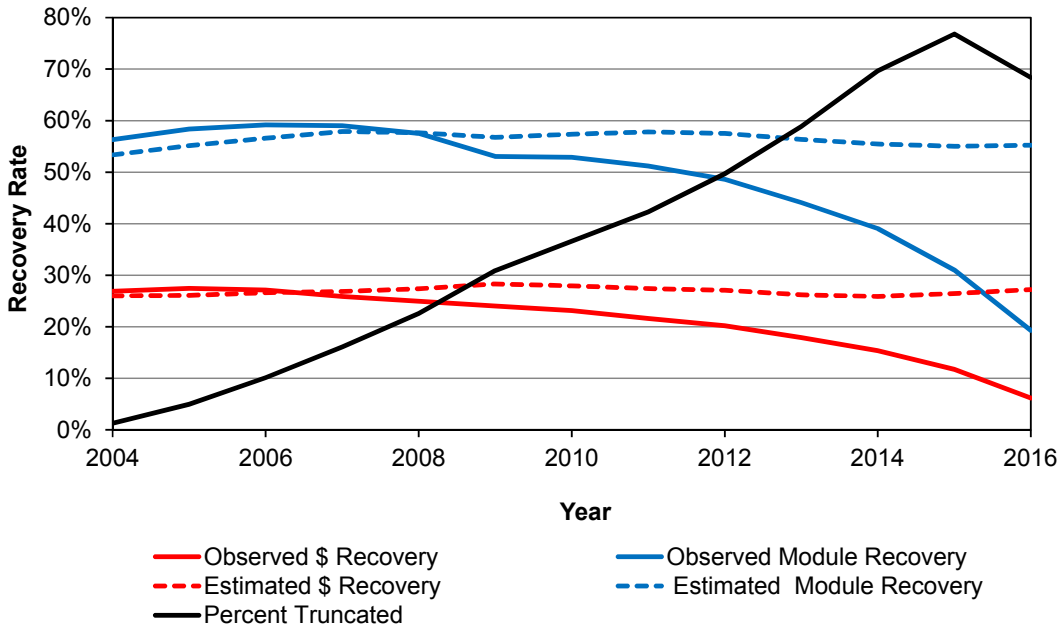
[8] U.S. Government Accountability Office (GAO) (2016, November), various years.

TABLE 1. Recovery Model Payment Truncation Control Parameter Estimates*

Form Number/Issue	Probability of Truncated Payments	Time Remaining at the End of the Sample
Form 940	2.732	-1.916
	(0.022)	(0.013)
Form 941	2.232	-1.125
	(0.011)	(0.003)
Form 1040	0.791	-0.776
	(0.009)	(0.002)
Form 1065	1.350	-1.327
	(0.025)	(0.012)
Form 1120	1.770	-0.999
	(0.035)	(0.007)
Trust Fund Recovery Penalty	-0.456	-0.601
	(0.015)	(0.003)
Business Other	1.330	-1.214
	(0.014)	(0.006)
Individual Other	0.198	-0.616
	(0.084)	(0.019)

NOTE: All estimates are significant at the 5% level. Standard errors reported in parentheses.
SOURCE: CFO Unpaid Assessments Inventory Recovery Model Output.

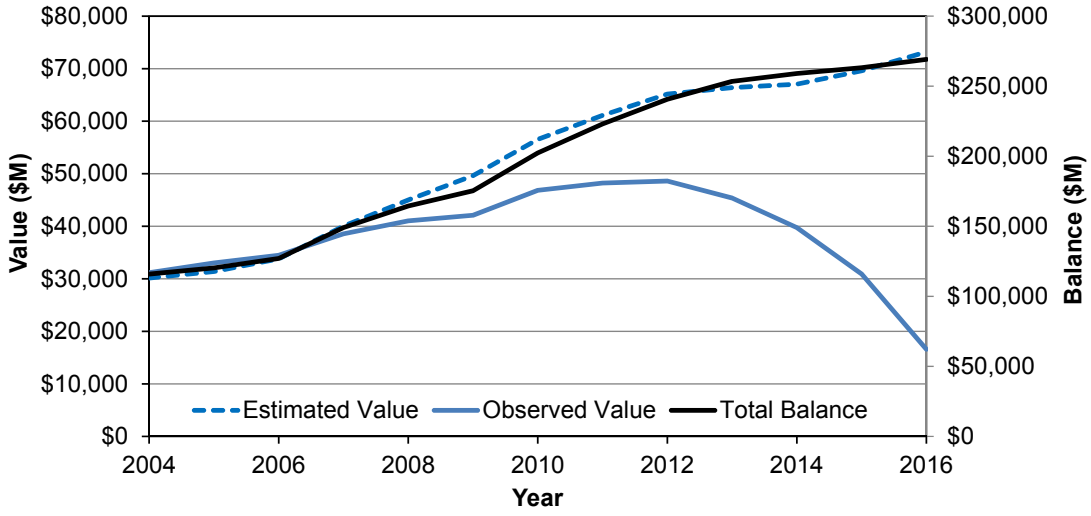
FIGURE 4. Individual Income Tax Form 1040—Average Percent of Modules and Dollars Recovered, Model Full-Statute Estimates and Actual Observed to Date



SOURCE: Form 1040 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

GAO audit results include net realizable values in the aggregate, whereas the recovery model approach can break out subpopulations by any form type. Figure 5 shows that the estimated value for Form 1040 assets tracks very closely to the total module balance for this population. The difference between the observed and estimated value is the truncation effect as stated before; the impact of the yet-to-be-made payments on the more recent assets over time not being currently observed. This is consistent with the recovery rates observed in Figure 2.

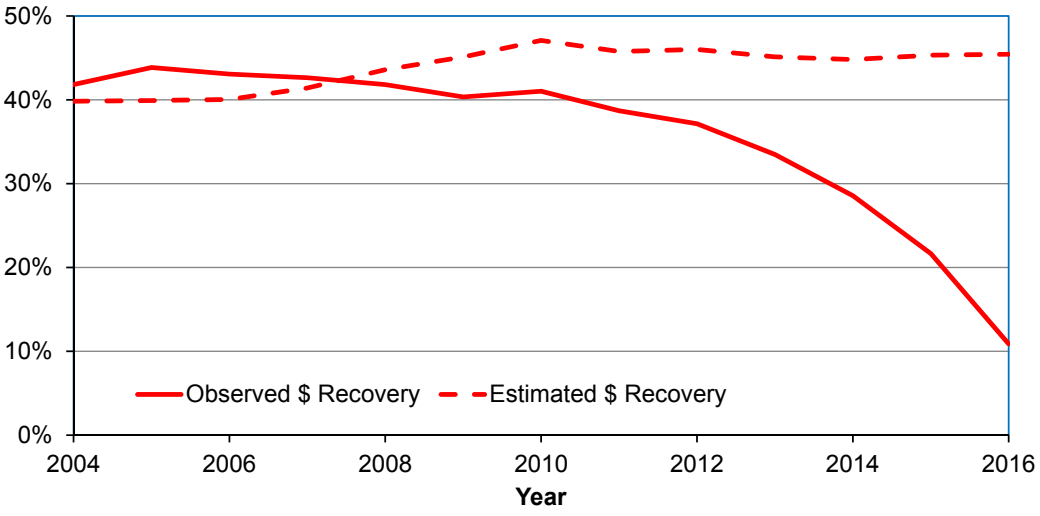
FIGURE 5. All Forms 1040—Aggregate Value and Module Balance



SOURCE: Form 1040 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

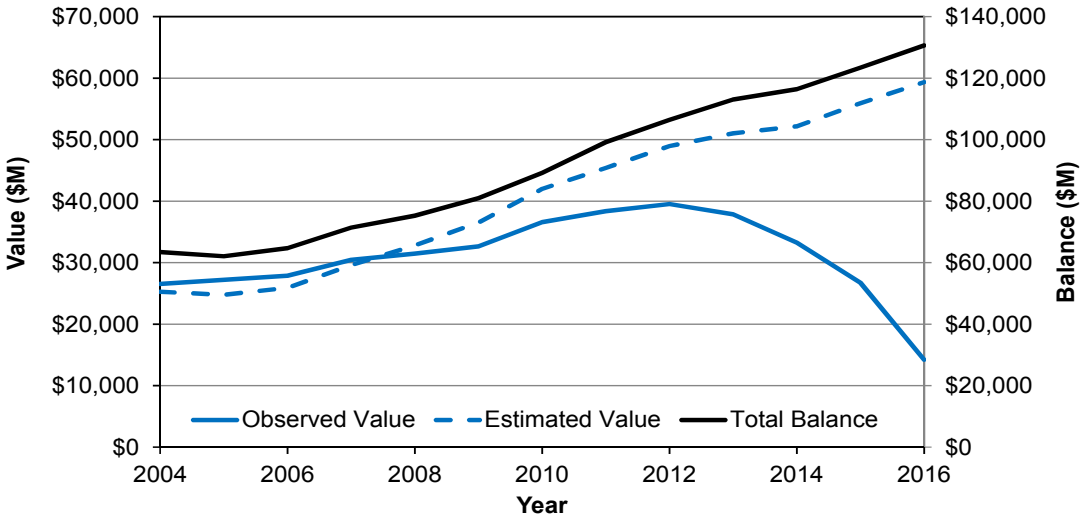
Figures 6 and 7 show the model estimates and observed actuals for the taxes receivable portion of Form 1040 UA inventory. These figures show higher average estimated recovery and actual observed rates than Form 1040 UA inventory. The taxes receivable recovery rates are higher than the total Form 1040 UA recovery rate because comparatively fewer payments are received on compliance assessments and write-offs. In Figure 7, the taxes receivable estimated value tracks the total module balance for Form 1040 taxes receivable.

FIGURE 6. Observed and Estimated Recovery Percentages of Form 1040 Modules—Taxes Receivable



SOURCE: Form 1040 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

FIGURE 7. Form 1040 Aggregate Value and Module Balance—Taxes Receivable

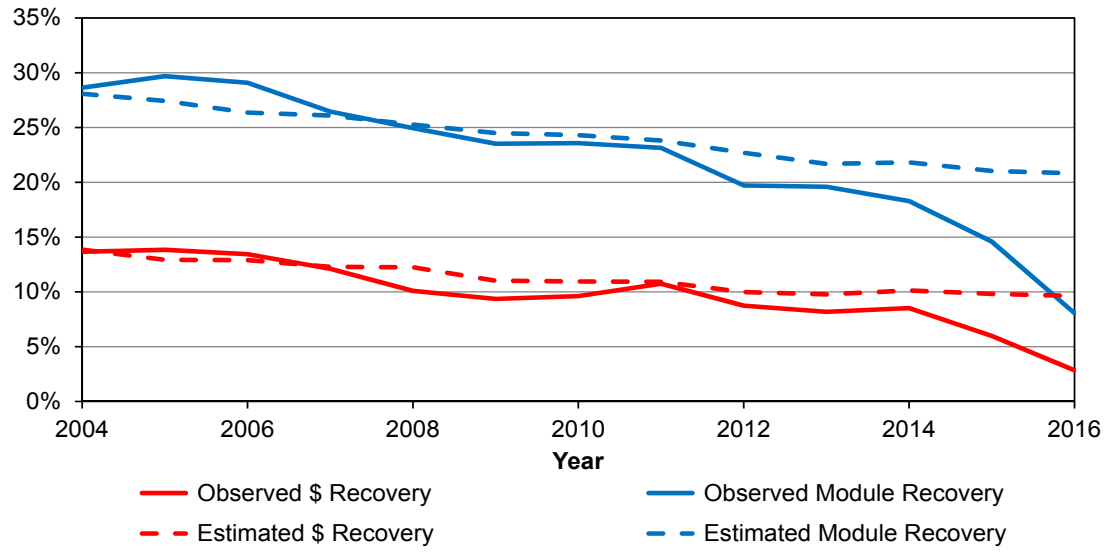


SOURCE: Form 1040 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

Figure 8 shows the model predicted average and actual observed recovery rates and recovered dollars for the business tax Form 941, *Employer's Quarterly Federal Tax Return,* UA inventory in Calendar Years January 2004 through 2016. Employers use this form to report income taxes and Social Security or Medicare taxes withheld from employees' paychecks, and to pay the employer's portion of Social Security or Medicare tax.

The recovery rates for the Form 941 UA inventory displayed in Figure 8 are lower and decrease at a faster rate than Form 1040 UA recovery rates. Actual observed dollars collected on Form 941 business tax UA is shown at nearly 14 percent, declining for more recent years as collections on more recent assets have not yet been realized. Like with Form 1040 UA, the difference between the estimated dollar recovery rate and the observed dollar recovery rate illustrates the truncation effect. The more rapid drop in recovery rates can be explained by the fact that business entities become "defunct" more frequently than individuals. As companies go out of business the payments on unpaid assessments may stop. As such, it takes fewer years to observe most of the payments, making the impact of truncated payments less pronounced for business tax UA inventory.

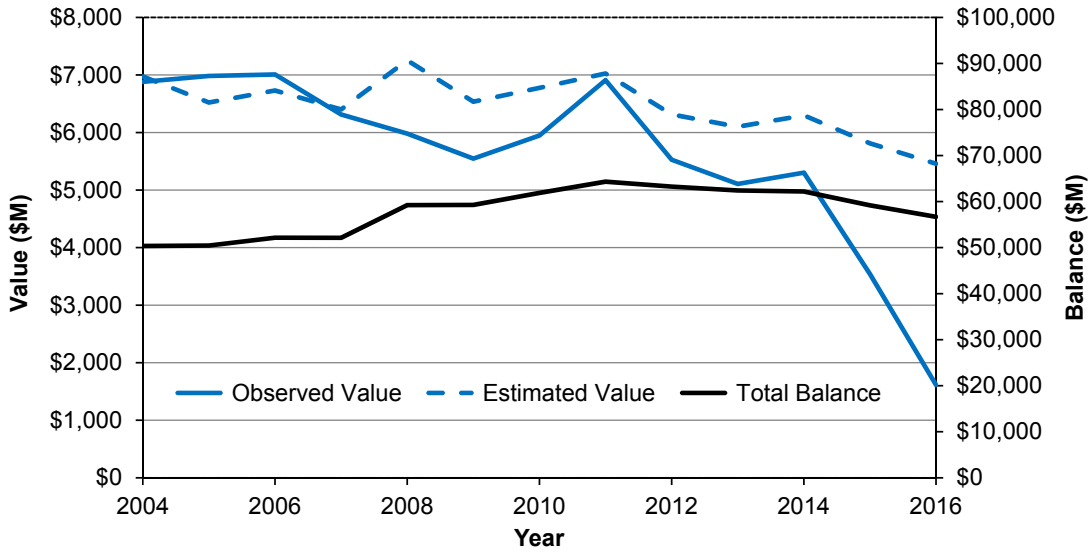
Figure 8. All Forms 941—Average Percentage of Modules and Dollars Recovered



SOURCE: Form 1040 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

In Figure 9, there is very little growth in the balance and a general decline in the value of the assets relative to the balance for Form 941 UA. The estimated value however does move with the actual value and declines.

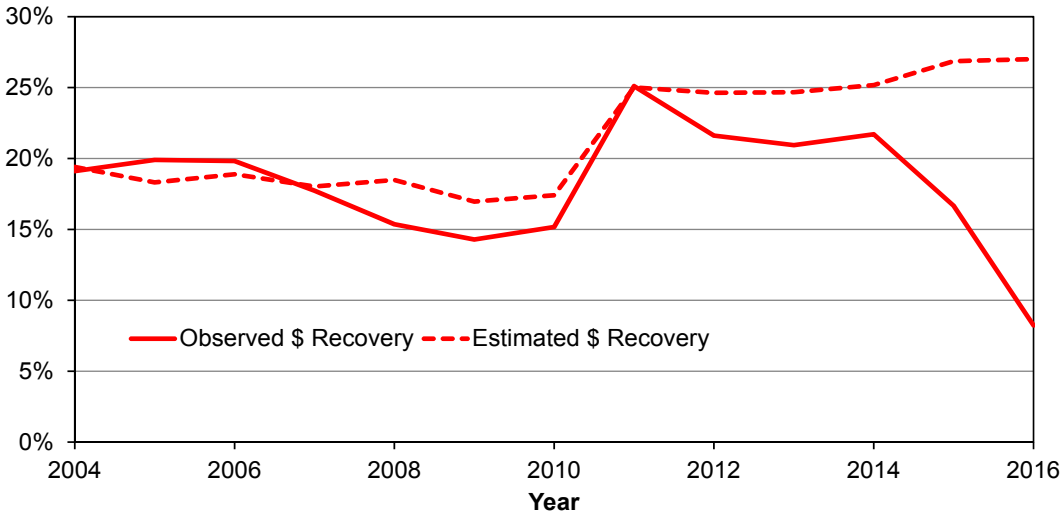
FIGURE 9. All Forms 941—Aggregate Value and Module Balance



SOURCE: Form 1040 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

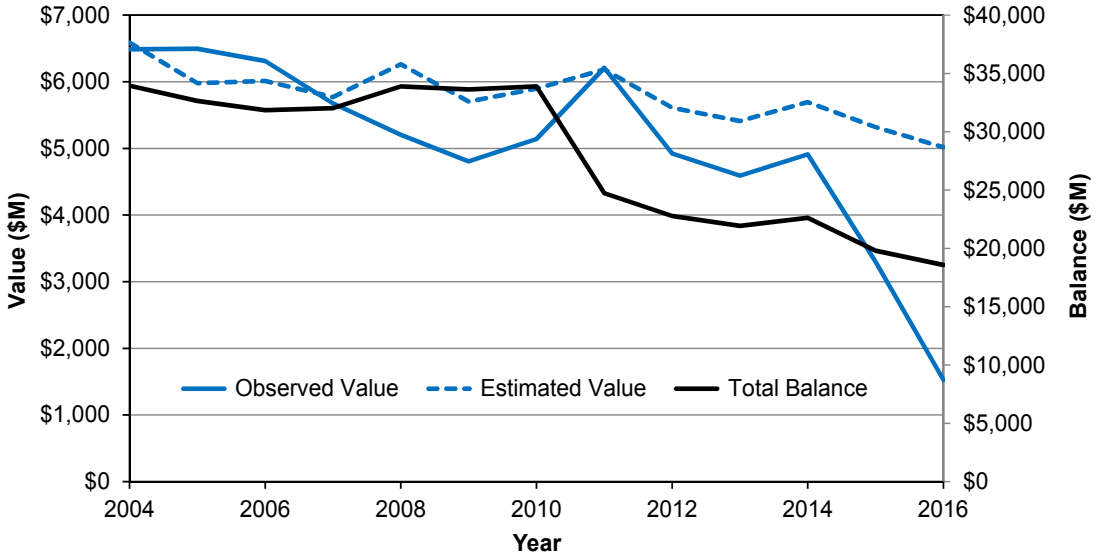
Figure 10 shows the dollar recovery rates for the taxes receivable portion of the Form 941 UA. From 2004 to 2016, the estimated recovery ranges from 19 to 26 percent. There is a noticeable jump in the rates, both estimated and observed for 2011. The value and total balance are reported in Figure 11. The estimated Form 941 UA value (black line) has fluctuated and declined from $6.5 billion to approximately $5 billion, and appears to move with the observed value. There is also a large decline in the aggregate balance (dashed line) from nearly $34 billion to $18.5 billion. This decline is proportionally larger than the decline in the value. There was a portion of the UA inventory that moved from taxes receivable to write-off. However, the model estimates were already heavily discounting these assets even before they moved to write-off. Thus, the estimated value does not decline.

FIGURE 10. Observed and Estimated Recovery Percentages of Form 941 Modules—Taxes Receivable



SOURCE: Form 941 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

FIGURE 11. Form 941 Aggregate Value and Module—Taxes Receivable



SOURCE: Form 941 Unpaid Assessments Inventory Recovery Model Results. Compliance Data Warehouse Accounts Receivable Dollar Inventory 2004–2016.

Conclusions

In this paper, we develop a predictive model of collectability for unpaid tax assessments. The primary role of this model is to provide an objective methodology for valuing UA that does not require time-intensive sampling and the inherent risks of financial misclassification. To predict the value of the inventory of unpaid assessments, we compile a repeated cross section of the UA inventory over multiple years. This allows us to capitalize on variation in the make-up of UA. However, it does introduce a unique truncation issue because of the pooling of the cross sections. We employ two statistical controls to account for this censoring.

The models will improve the objectivity of financial management reporting and eliminate sampling errors, thereby improving the precision of the IRS financial statement audit. The entire inventory of UA can be continuously rescored and valued, with little resource costs. Explanations for changes in the aggregate value of UA can easily be traced back to changes in the inventory and the associated case characteristics.

References

Chief Financial Officers (CFO) Act of 1990 (1990, November). Public Law 101–576, 101st Congress. Retrieved from http://www.gao.gov/special.pubs/af12194.pdf.

Civic Impulse (2017). S. 2170—103rd Congress: Government Management Reform Act of 1994. Retrieved from https://www.govtrack.us/congress/bills/103/s2170.

Federal Accounting Standards Advisory Board (FASAB), Accounting for Revenue and Other Financing Sources and Concepts for Reconciling Budgetary and Financial Accounting, Statement of Recommended Accounting Standards Number 7. Retrieved from http://www.fasab.gov/pdffiles/sffas-7.pdf.

Financial Sector Advisory Centre (FINSAC) (1994, August). Retrieved from http://pubdocs.worldbank.org/en/314911450690270267/FinSAC-LoanClassification-Provisioning-Paper.pdf.

Global Public Policy Committee (GPPC) (2016, June). The Implementation of IFRS 9 Impairment Requirements by Banks. Considerations for Those Charged with Governance of Systemically Important Banks. Global Public Policy Committee of Representatives of the Six Largest Accounting Networks. Retrieved from http://www.ey.com/Publication/vwLUAssets/Implementation_of_IFRS_9_impairment_requirements_by_systemically_important_banks/$FILE/BCM-FIImpair-GPPC-June2016%20int.pdf, 2.8.2.1.

Miller, Erik, Stacy Orlett, and Alex Turk (2014). "Uncollectible Versus Unproductive: Compliance Impact of Working Collection Cases That Are Ultimately Not Fully Collectible," *IRS Research Bulletin: Papers Given at the 2014 IRS-Tax Policy Center Research Conference,* pp. 79–98.

Moody's Investor Service, Global Credit Research (2006, August). Retrieved from http://care-mendoza.nd.edu/assets/152347/loss-given-default-rating-methodology.pdf.

Turk, Alex, and Terry Ashley (2002). "Accounts Receivable Resolution and the Impact of Lien Filing Policy on Sole Proprietor Businesses," *12th Annual Federal Forecasters Conference 2002—Papers and Proceedings*, pp. 323–331.

Turk, Alex, John Iuranich, Stacy Orlett, and Saurabh Datta (2016). "Resolving Unpaid Taxes and the Notice of Federal Tax Lien: Evidence from the Fresh Start Initiative," *IRS Research Bulletin: Papers Given at the 6th Annual Joint Research Conference on Tax Administration,* pp. 13–29.

U.S. Government Accountability Office (GAO) (2016, November). Financial Audit: Examination of IRS' 2016 and 2015 Fiscal Year Financial Statements, GAO-17-140.

_____ (2015, November). Financial Audit: Examination of IRS' 2015 and 2014 Fiscal Year Financial Statements, GAO-16-140.

_____ (2014, November). Financial Audit: Examination of IRS' 2014 and 2013 Fiscal Year Financial Statements, GAO-15-143.

_____ (2013, December). Financial Audit: Examination of IRS' 2013 and 2012 Fiscal Year Financial Statements, GAO-14-169.

_____ (2012, November). Financial Audit: Examination of IRS' 2012 and 2011 Fiscal Year Financial Statements, GAO-13-120.

_____ (2011, November). Financial Audit: Examination of IRS' 2011 and 2010 Fiscal Year Financial Statements, GAO-12-165.

_____ (2010, November). Financial Audit: Examination of IRS' 2010 and 2009 Fiscal Year Financial Statements, GAO-11-142.

_____ (2009, November). Financial Audit: Examination of IRS' 2009 and 2008 Fiscal Year Financial Statements, GAO-10-176.

Appendix

TABLE A1. Form 941 Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	38%	38%	38%	1%	$33,950	$6,486	$6,588
	2005	41%	39%	38%	1%	$32,640	$6,495	$5,977
	2006	44%	41%	41%	3%	$31,839	$6,311	$6,013
	2007	40%	43%	40%	9%	$32,016	$5,677	$5,771
	2008	40%	45%	42%	20%	$33,878	$5,204	$6,262
	2009	39%	43%	41%	30%	$33,633	$4,804	$5,704
	2010	40%	41%	42%	37%	$33,884	$5,140	$5,898
	2011	45%	43%	47%	43%	$24,728	$6,208	$6,184
	2012	39%	41%	46%	48%	$22,775	$4,923	$5,610
	2013	41%	38%	46%	54%	$21,924	$4,591	$5,409
	2014	39%	35%	46%	59%	$22,626	$4,911	$5,695
	2015	32%	27%	45%	68%	$19,798	$3,300	$5,318
	2016	18%	15%	45%	83%	$18,576	$1,529	$5,017
Compliance Assessments	2004	15%	11%	11%	0%	$4,084	$293	$295
	2005	17%	11%	10%	1%	$4,624	$308	$377
	2006	22%	16%	16%	3%	$4,916	$354	$462
	2007	19%	17%	15%	17%	$4,757	$391	$383
	2008	19%	19%	16%	36%	$7,661	$493	$698
	2009	16%	17%	14%	50%	$6,716	$494	$519
	2010	15%	15%	13%	55%	$7,259	$526	$568
	2011	14%	14%	12%	62%	$5,766	$421	$497
	2012	10%	12%	10%	67%	$5,258	$342	$377
	2013	10%	12%	10%	74%	$5,067	$290	$380
	2014	9%	10%	10%	79%	$4,911	$221	$319
	2015	7%	7%	9%	86%	$4,281	$119	$227
	2016	3%	3%	9%	94%	$3,883	$53	$194
Write-off	2004	2%	2%	2%	0%	$12,315	$101	$93
	2005	5%	4%	4%	0%	$13,193	$181	$167
	2006	7%	5%	5%	0%	$15,398	$344	$255
	2007	5%	5%	4%	1%	$15,352	$245	$248
	2008	5%	6%	4%	6%	$17,692	$284	$296
	2009	4%	6%	4%	13%	$18,931	$249	$312
	2010	4%	5%	4%	22%	$20,735	$282	$311
	2011	3%	4%	3%	34%	$33,831	$282	$346
	2012	3%	3%	3%	46%	$35,210	$262	$328
	2013	2%	2%	3%	57%	$35,433	$226	$312
	2014	2%	2%	2%	69%	$34,653	$170	$283
	2015	1%	1%	2%	81%	$35,107	$122	$268
	2016	0%	1%	2%	94%	$34,220	$31	$247
All	2004	29%	28%	28%	1%	$50,349	$6,880	$6,975
	2005	30%	28%	27%	1%	$50,457	$6,984	$6,521
	2006	29%	27%	26%	2%	$52,154	$7,009	$6,729
	2007	26%	28%	26%	7%	$52,125	$6,312	$6,402
	2008	25%	28%	25%	16%	$59,231	$5,980	$7,256
	2009	24%	26%	24%	25%	$59,280	$5,548	$6,535
	2010	24%	24%	24%	34%	$61,878	$5,948	$6,777
	2011	23%	23%	24%	41%	$64,326	$6,911	$7,026
	2012	20%	20%	23%	49%	$63,244	$5,527	$6,315
	2013	20%	18%	22%	57%	$62,424	$5,107	$6,101
	2014	18%	17%	22%	66%	$62,190	$5,302	$6,298
	2015	15%	13%	21%	76%	$59,185	$3,541	$5,813
	2016	8%	7%	21%	90%	$56,679	$1,613	$5,458

TABLE A2. Form 1040 Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	70%	66%	66%	2%	$63,434	$26,527	$25,267
	2005	72%	68%	68%	5%	$62,073	$27,220	$24,770
	2006	71%	70%	68%	9%	$64,730	$27,874	$25,918
	2007	71%	71%	70%	14%	$71,421	$30,446	$29,576
	2008	70%	72%	70%	18%	$75,256	$31,470	$32,800
	2009	65%	70%	69%	25%	$80,944	$32,655	$36,516
	2010	65%	68%	70%	29%	$89,183	$36,590	$41,989
	2011	63%	65%	70%	35%	$99,170	$38,376	$45,388
	2012	60%	60%	70%	41%	$106,428	$39,527	$48,964
	2013	55%	53%	69%	50%	$113,063	$37,859	$51,022
	2014	49%	46%	69%	62%	$116,404	$33,251	$52,154
	2015	39%	35%	68%	71%	$123,411	$26,737	$55,925
	2016	24%	20%	68%	62%	$130,613	$14,229	$59,345
Compliance Assessments	2004	50%	48%	48%	1%	$16,156	$2,954	$3,272
	2005	56%	55%	54%	11%	$20,139	$3,841	$4,814
	2006	55%	57%	54%	23%	$28,880	$4,710	$6,070
	2007	53%	57%	54%	32%	$41,646	$6,066	$8,438
	2008	49%	53%	51%	40%	$44,639	$7,284	$9,838
	2009	46%	50%	51%	47%	$47,950	$7,321	$10,686
	2010	46%	47%	52%	49%	$53,619	$7,119	$11,053
	2011	42%	41%	49%	56%	$77,558	$8,093	$13,594
	2012	39%	37%	49%	62%	$80,045	$7,400	$13,927
	2013	35%	33%	49%	68%	$69,407	$5,970	$12,615
	2014	34%	30%	50%	74%	$66,122	$5,307	$12,033
	2015	28%	24%	49%	78%	$56,184	$3,452	$10,868
	2016	20%	16%	50%	62%	$55,522	$2,022	$11,067
Write-off	2004	18%	16%	16%	0%	$36,299	$1,690	$1,612
	2005	19%	17%	16%	1%	$38,059	$1,959	$1,796
	2006	20%	18%	17%	4%	$33,439	$1,901	$1,805
	2007	20%	19%	17%	12%	$35,996	$2,054	$2,006
	2008	20%	19%	17%	22%	$44,584	$2,277	$2,363
	2009	16%	18%	16%	35%	$46,433	$2,121	$2,429
	2010	17%	18%	16%	51%	$59,527	$3,134	$3,489
	2011	16%	17%	17%	55%	$46,288	$1,753	$2,139
	2012	14%	15%	16%	69%	$54,205	$1,681	$2,286
	2013	11%	12%	16%	82%	$70,989	$1,547	$2,755
	2014	8%	9%	15%	94%	$76,506	$1,182	$2,853
	2015	6%	6%	15%	96%	$83,644	$712	$2,801
	2016	3%	4%	15%	98%	$82,995	$336	$2,777
All	2004	56%	53%	53%	1%	$115,888	$31,171	$30,150
	2005	58%	56%	55%	5%	$120,270	$33,020	$31,380
	2006	59%	58%	57%	10%	$127,049	$34,486	$33,793
	2007	59%	60%	58%	16%	$149,062	$38,565	$40,020
	2008	58%	59%	58%	23%	$164,478	$41,031	$45,001
	2009	53%	57%	57%	31%	$175,327	$42,097	$49,631
	2010	53%	55%	57%	37%	$202,329	$46,843	$56,531
	2011	51%	53%	58%	42%	$223,016	$48,223	$61,122
	2012	49%	48%	58%	50%	$240,678	$48,608	$65,177
	2013	44%	43%	56%	59%	$253,459	$45,376	$66,392
	2014	39%	37%	55%	70%	$259,033	$39,740	$67,040
	2015	31%	28%	55%	77%	$263,239	$30,901	$69,595
	2016	19%	16%	55%	68%	$269,129	$16,586	$73,189

TABLE A3. Form 940 Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	31%	31%	31%	1%	$609	$123	$117
	2005	34%	33%	32%	1%	$907	$125	$170
	2006	38%	35%	34%	3%	$613	$116	$118
	2007	34%	37%	34%	8%	$802	$113	$149
	2008	34%	39%	36%	21%	$755	$108	$128
	2009	32%	35%	35%	33%	$754	$85	$118
	2010	33%	33%	36%	42%	$695	$85	$105
	2011	33%	31%	36%	49%	$559	$87	$107
	2012	29%	29%	35%	57%	$581	$78	$113
	2013	29%	27%	36%	64%	$551	$78	$105
	2014	26%	24%	35%	72%	$566	$79	$111
	2015	20%	18%	35%	78%	$505	$71	$115
	2016	9%	8%	35%	91%	$507	$29	$114
Compliance Assessments	2004	17%	13%	13%	1%	$948	$80	$83
	2005	19%	14%	14%	1%	$864	$90	$77
	2006	20%	16%	15%	2%	$880	$78	$82
	2007	17%	16%	14%	11%	$973	$74	$89
	2008	18%	20%	16%	28%	$1,003	$79	$99
	2009	15%	17%	14%	45%	$1,050	$61	$106
	2010	16%	16%	13%	57%	$822	$73	$70
	2011	13%	14%	12%	65%	$817	$60	$67
	2012	10%	13%	11%	72%	$902	$44	$84
	2013	8%	9%	8%	81%	$689	$50	$44
	2014	7%	9%	8%	85%	$692	$31	$49
	2015	7%	8%	10%	88%	$628	$26	$55
	2016	3%	3%	9%	96%	$649	$8	$49
Write-off	2004	1%	1%	1%	0%	$2,151	$9	$11
	2005	3%	2%	2%	0%	$1,907	$12	$13
	2006	4%	3%	2%	0%	$1,970	$16	$15
	2007	3%	3%	2%	1%	$1,832	$19	$15
	2008	3%	3%	2%	5%	$2,101	$24	$16
	2009	2%	3%	2%	12%	$2,214	$21	$17
	2010	2%	3%	2%	20%	$2,319	$17	$17
	2011	2%	2%	2%	30%	$2,489	$12	$16
	2012	2%	2%	2%	40%	$2,454	$12	$15
	2013	2%	2%	2%	51%	$2,436	$10	$14
	2014	1%	1%	2%	62%	$2,425	$8	$12
	2015	1%	1%	2%	77%	$2,436	$4	$12
	2016	0%	0%	2%	91%	$2,379	$1	$11
All	2004	18%	17%	17%	0%	$3,708	$213	$212
	2005	19%	17%	17%	1%	$3,677	$227	$260
	2006	20%	18%	18%	2%	$3,462	$210	$216
	2007	19%	20%	18%	6%	$3,607	$206	$252
	2008	18%	20%	18%	15%	$3,859	$210	$242
	2009	16%	18%	17%	25%	$4,018	$167	$241
	2010	17%	17%	17%	35%	$3,836	$176	$191
	2011	16%	16%	17%	43%	$3,865	$159	$190
	2012	14%	14%	17%	52%	$3,937	$133	$211
	2013	14%	13%	17%	60%	$3,677	$138	$163
	2014	12%	11%	16%	69%	$3,682	$118	$171
	2015	9%	8%	16%	79%	$3,569	$101	$182
	2016	4%	3%	15%	91%	$3,535	$38	$174

TABLE A4. Form 1065 Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	18%	18%	18%	1%	$133	$14	$15
	2005	20%	18%	18%	0%	$101	$12	$11
	2006	34%	30%	30%	0%	$111	$24	$22
	2007	29%	30%	29%	11%	$138	$26	$27
	2008	29%	31%	30%	24%	$144	$29	$30
	2009	23%	27%	28%	34%	$220	$33	$43
	2010	24%	28%	27%	41%	$282	$50	$60
	2011	24%	22%	26%	49%	$321	$59	$67
	2012	18%	18%	23%	56%	$465	$90	$86
	2013	18%	15%	21%	64%	$549	$72	$95
	2014	15%	13%	20%	67%	$612	$68	$101
	2015	12%	10%	19%	73%	$680	$61	$102
	2016	6%	6%	19%	85%	$726	$33	$105
Compliance Assessments	2004	19%	19%	19%	27%	$1	$0	$0
	2005	28%	29%	29%	3%	$1	$0	$0
	2006	30%	25%	25%	4%	$105	$1	$2
	2007	36%	25%	22%	4%	$10	$1	$1
	2008	22%	28%	26%	13%	$1	$1	$1
	2009	12%	21%	19%	25%	$2	$1	$1
	2010	41%	38%	35%	32%	$27	$27	$7
	2011	16%	15%	25%	57%	$130	$0	$3
	2012	30%	42%	36%	36%	$134	$0	$2
	2013	30%	39%	29%	30%	$0	$0	$0
	2014	8%	21%	25%	54%	$12	$2	$1
	2015	23%	19%	26%	67%	$3	$3	$0
	2016	50%	37%	30%	50%	$46	$14	$6
Write-off	2004	3%	2%	2%	1%	$44	$1	$0
	2005	5%	3%	3%	1%	$61	$1	$1
	2006	5%	4%	4%	1%	$82	$2	$1
	2007	4%	5%	4%	4%	$78	$1	$1
	2008	4%	5%	4%	11%	$73	$2	$2
	2009	4%	5%	5%	24%	$69	$2	$2
	2010	4%	6%	5%	37%	$74	$2	$2
	2011	4%	4%	5%	52%	$103	$2	$3
	2012	3%	3%	4%	65%	$141	$3	$4
	2013	3%	2%	4%	73%	$197	$4	$5
	2014	2%	1%	3%	81%	$282	$3	$7
	2015	1%	1%	3%	87%	$398	$2	$9
	2016	0%	0%	2%	98%	$483	$1	$10
All	2004	15%	15%	15%	1%	$177	$15	$15
	2005	15%	13%	13%	1%	$163	$14	$12
	2006	24%	21%	21%	1%	$298	$27	$25
	2007	23%	24%	23%	9%	$225	$28	$29
	2008	22%	24%	23%	20%	$218	$32	$32
	2009	18%	21%	22%	32%	$291	$36	$46
	2010	19%	23%	22%	40%	$383	$79	$68
	2011	19%	17%	20%	50%	$554	$62	$74
	2012	13%	13%	17%	59%	$740	$93	$92
	2013	12%	10%	15%	67%	$746	$76	$101
	2014	9%	8%	13%	73%	$906	$73	$109
	2015	7%	5%	11%	80%	$1,081	$66	$111
	2016	3%	3%	11%	91%	$1,255	$48	$121

TABLE A5. Form 1120 Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	40%	40%	40%	1%	$1,232	$431	$372
	2005	43%	41%	41%	1%	$703	$190	$161
	2006	47%	44%	44%	1%	$1,474	$634	$463
	2007	46%	47%	45%	7%	$3,259	$1,694	$1,086
	2008	49%	50%	49%	16%	$986	$464	$346
	2009	46%	47%	48%	26%	$1,716	$924	$728
	2010	48%	54%	52%	31%	$1,685	$642	$616
	2011	44%	42%	49%	41%	$1,822	$403	$625
	2012	33%	36%	46%	51%	$1,807	$210	$500
	2013	32%	30%	43%	60%	$2,922	$227	$460
	2014	27%	25%	41%	66%	$1,841	$435	$484
	2015	22%	19%	39%	73%	$1,998	$512	$486
	2016	12%	11%	38%	86%	$1,520	$145	$487
Compliance Assessments	2004	12%	14%	14%	7%	$1,752	$136	$29
	2005	17%	16%	15%	7%	$1,675	$10	$9
	2006	25%	17%	16%	10%	$1,947	$2	$35
	2007	23%	20%	17%	17%	$2,323	$4	$70
	2008	20%	23%	18%	36%	$3,440	$14	$139
	2009	25%	23%	18%	39%	$2,620	$88	$47
	2010	20%	23%	19%	54%	$3,047	$13	$38
	2011	13%	14%	17%	69%	$3,891	$152	$113
	2012	11%	15%	17%	70%	$3,865	$27	$99
	2013	14%	12%	16%	76%	$3,324	$3	$52
	2014	14%	14%	17%	75%	$3,186	$9	$37
	2015	9%	11%	17%	84%	$3,193	$1	$16
	2016	4%	5%	14%	92%	$3,438	$2	$27
Write-off	2004	1%	1%	1%	1%	$13,533	$253	$15
	2005	3%	3%	3%	2%	$9,447	$259	$11
	2006	3%	3%	3%	1%	$9,552	$279	$10
	2007	3%	3%	3%	2%	$9,308	$26	$11
	2008	2%	3%	3%	5%	$9,312	$4	$14
	2009	2%	3%	3%	10%	$9,629	$5	$12
	2010	3%	5%	3%	20%	$10,128	$6	$13
	2011	3%	4%	4%	36%	$10,602	$8	$11
	2012	3%	3%	4%	56%	$11,098	$18	$13
	2013	3%	2%	4%	72%	$11,853	$3	$14
	2014	2%	1%	4%	85%	$2,987	$3	$18
	2015	1%	1%	4%	93%	$3,200	$2	$15
	2016	1%	0%	3%	98%	$3,897	$1	$23
All	2004	32%	32%	31%	1%	$16,517	$820	$415
	2005	32%	30%	30%	1%	$11,824	$460	$181
	2006	31%	29%	28%	1%	$12,974	$915	$507
	2007	32%	32%	31%	5%	$14,890	$1,724	$1,166
	2008	32%	33%	32%	12%	$13,738	$481	$499
	2009	30%	31%	32%	20%	$13,965	$1,017	$787
	2010	37%	42%	40%	29%	$14,860	$660	$667
	2011	35%	34%	39%	40%	$16,315	$563	$750
	2012	26%	28%	37%	52%	$16,770	$256	$611
	2013	24%	23%	33%	63%	$18,099	$233	$526
	2014	20%	18%	30%	71%	$8,014	$447	$538
	2015	15%	13%	27%	80%	$8,391	$515	$517
	2016	8%	7%	25%	90%	$8,856	$147	$536

TABLE A6. TFRP Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	61%	49%	49%	5%	$59	$15	$22
	2005	59%	46%	47%	10%	$54	$15	$18
	2006	56%	43%	45%	16%	$49	$14	$19
	2007	34%	42%	46%	41%	$2,806	$455	$1,185
	2008	30%	36%	42%	45%	$4,649	$636	$1,646
	2009	27%	31%	40%	52%	$6,383	$738	$2,050
	2010	26%	28%	40%	58%	$6,776	$791	$2,271
	2011	25%	25%	40%	66%	$7,799	$861	$2,644
	2012	24%	23%	41%	70%	$8,590	$927	$2,889
	2013	22%	19%	41%	77%	$9,333	$910	$3,116
	2014	18%	14%	39%	88%	$9,698	$740	$2,940
	2015	14%	10%	37%	91%	$9,692	$560	$2,850
	2016	8%	6%	36%	87%	$9,403	$287	$2,691
Compliance Assessments	2004	28%	32%	33%	9%	$12,598	$1,411	$2,729
	2005	28%	31%	33%	20%	$12,609	$1,471	$2,814
	2006	28%	30%	32%	28%	$12,617	$1,458	$2,847
	2007	23%	25%	28%	30%	$9,872	$883	$1,874
	2008	22%	24%	29%	41%	$7,592	$583	$1,477
	2009	21%	22%	31%	51%	$5,988	$504	$1,354
	2010	22%	20%	32%	56%	$5,530	$455	$1,267
	2011	21%	18%	33%	62%	$5,484	$480	$1,436
	2012	19%	16%	33%	69%	$5,906	$485	$1,570
	2013	16%	12%	32%	80%	$5,400	$380	$1,412
	2014	14%	9%	32%	91%	$4,949	$305	$1,314
	2015	10%	6%	30%	94%	$4,539	$178	$1,147
	2016	5%	3%	30%	66%	$4,321	$74	$1,131
Write-off	2004	22%	19%	19%	2%	$47	$2	$4
	2005	28%	19%	20%	5%	$39	$2	$6
	2006	24%	17%	18%	9%	$38	$2	$6
	2007	19%	18%	21%	42%	$398	$35	$88
	2008	15%	14%	17%	45%	$1,296	$58	$164
	2009	10%	11%	15%	51%	$1,858	$56	$187
	2010	9%	9%	15%	57%	$2,617	$70	$279
	2011	7%	8%	12%	53%	$3,132	$56	$260
	2012	6%	6%	11%	66%	$3,309	$55	$269
	2013	5%	5%	11%	82%	$3,797	$56	$322
	2014	4%	4%	11%	98%	$3,687	$46	$294
	2015	3%	3%	10%	98%	$3,661	$26	$299
	2016	1%	2%	10%	99%	$3,984	$12	$324
All	2004	30%	33%	34%	9%	$12,705	$1,428	$2,755
	2005	30%	32%	33%	19%	$12,702	$1,489	$2,839
	2006	29%	30%	33%	27%	$12,704	$1,474	$2,871
	2007	27%	31%	34%	35%	$13,077	$1,373	$3,147
	2008	25%	29%	34%	43%	$13,536	$1,277	$3,288
	2009	23%	26%	34%	51%	$14,230	$1,298	$3,591
	2010	22%	23%	33%	57%	$14,924	$1,317	$3,816
	2011	21%	20%	34%	63%	$16,416	$1,396	$4,340
	2012	20%	18%	34%	69%	$17,805	$1,468	$4,728
	2013	17%	15%	33%	78%	$18,530	$1,347	$4,851
	2014	14%	11%	32%	90%	$18,335	$1,091	$4,548
	2015	11%	8%	31%	93%	$17,891	$764	$4,296
	2016	6%	4%	29%	84%	$17,709	$373	$4,145

TABLE A7. Other Individual Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	31%	37%	47%	0%	$440	$39	$197
	2005	49%	49%	49%	2%	$547	$58	$322
	2006	48%	48%	47%	9%	$1,328	$288	$627
	2007	51%	51%	47%	15%	$2,455	$648	$1,046
	2008	49%	48%	47%	22%	$1,919	$365	$789
	2009	48%	48%	48%	27%	$2,125	$409	$883
	2010	47%	48%	48%	35%	$2,554	$435	$1,097
	2011	47%	48%	47%	44%	$2,683	$472	$1,250
	2012	46%	46%	46%	50%	$2,982	$544	$1,321
	2013	44%	44%	47%	59%	$3,534	$572	$1,672
	2014	40%	40%	46%	73%	$4,948	$1,005	$2,242
	2015	32%	39%	47%	80%	$5,881	$968	$2,730
	2016	23%	57%	49%	25%	$5,413	$235	$2,560
Compliance Assessments	2004	39%	32%	53%	4%	$67	$4	$44
	2005	29%	22%	49%	3%	$186	$6	$79
	2006	39%	37%	53%	12%	$406	$40	$239
	2007	39%	42%	48%	18%	$408	$41	$228
	2008	49%	45%	48%	17%	$197	$33	$123
	2009	43%	45%	45%	28%	$253	$26	$146
	2010	49%	44%	44%	39%	$246	$29	$120
	2011	44%	44%	45%	54%	$314	$31	$126
	2012	35%	41%	43%	60%	$447	$23	$220
	2013	42%	39%	44%	66%	$432	$122	$125
	2014	34%	38%	51%	75%	$214	$25	$125
	2015	27%	34%	48%	82%	$249	$24	$116
	2016	13%	35%	46%	62%	$297	$15	$110
Write-off	2004	15%	11%	49%	3%	$259	$14	$121
	2005	13%	12%	50%	3%	$443	$12	$191
	2006	16%	15%	49%	6%	$729	$18	$435
	2007	15%	15%	49%	6%	$767	$12	$455
	2008	15%	15%	48%	12%	$976	$22	$483
	2009	13%	13%	50%	18%	$754	$25	$366
	2010	12%	13%	48%	23%	$807	$21	$400
	2011	12%	12%	48%	29%	$769	$18	$384
	2012	11%	12%	46%	45%	$763	$16	$335
	2013	11%	12%	48%	72%	$1,005	$24	$470
	2014	8%	9%	46%	95%	$1,165	$20	$564
	2015	5%	10%	47%	97%	$1,249	$11	$552
	2016	3%	10%	46%	96%	$1,566	$5	$640
All	2004	25%	26%	48%	1%	$765	$57	$362
	2005	35%	34%	49%	2%	$1,176	$76	$591
	2006	40%	39%	48%	8%	$2,463	$347	$1,300
	2007	44%	44%	48%	13%	$3,631	$701	$1,730
	2008	40%	39%	48%	19%	$3,092	$420	$1,396
	2009	39%	39%	48%	25%	$3,132	$459	$1,394
	2010	40%	40%	48%	33%	$3,607	$485	$1,616
	2011	40%	40%	47%	41%	$3,766	$521	$1,759
	2012	39%	40%	46%	50%	$4,193	$582	$1,875
	2013	37%	37%	47%	62%	$4,971	$718	$2,267
	2014	34%	34%	46%	77%	$6,328	$1,050	$2,931
	2015	26%	33%	47%	83%	$7,379	$1,004	$3,398
	2016	21%	52%	49%	32%	$7,277	$254	$3,313

TABLE A8. Other Business Recovery Model Results ($ in Millions)

Financial Class	Year	Actual Recovery	Predicted Recovery with Truncation	Predicted Recovery	Percent Truncated	Total Balance	Observed Value	Estimated Value
Taxes Receivable	2004	48%	47%	47%	1%	$7,492	$4,862	$2,338
	2005	50%	47%	47%	1%	$7,037	$4,456	$2,114
	2006	44%	43%	42%	1%	$7,140	$4,612	$1,940
	2007	41%	42%	42%	5%	$8,458	$5,401	$2,353
	2008	41%	43%	43%	12%	$8,618	$5,544	$2,587
	2009	39%	42%	43%	20%	$9,488	$6,110	$2,908
	2010	38%	39%	43%	27%	$10,800	$6,942	$3,212
	2011	38%	36%	43%	35%	$9,810	$5,500	$2,780
	2012	33%	35%	43%	40%	$10,085	$5,088	$2,496
	2013	34%	32%	43%	49%	$8,754	$3,823	$2,209
	2014	31%	32%	44%	53%	$8,785	$2,908	$2,260
	2015	28%	26%	44%	61%	$8,355	$2,061	$2,079
	2016	19%	16%	45%	77%	$8,426	$1,394	$2,165
Compliance Assessments	2004	19%	19%	18%	0%	$8,690	$161	$587
	2005	20%	19%	18%	0%	$8,927	$239	$813
	2006	21%	20%	19%	1%	$4,179	$202	$568
	2007	19%	21%	19%	12%	$2,940	$191	$397
	2008	19%	21%	20%	23%	$4,107	$179	$429
	2009	19%	19%	21%	36%	$2,715	$211	$415
	2010	18%	17%	20%	45%	$2,454	$207	$380
	2011	15%	15%	20%	54%	$2,471	$219	$366
	2012	13%	13%	21%	64%	$2,372	$159	$389
	2013	12%	11%	20%	75%	$2,802	$143	$439
	2014	11%	10%	20%	78%	$2,318	$136	$346
	2015	9%	6%	19%	85%	$2,262	$132	$323
	2016	5%	3%	19%	94%	$1,938	$59	$301
Write-off	2004	3%	3%	3%	1%	$2,894	$53	$34
	2005	6%	5%	5%	1%	$3,500	$65	$52
	2006	9%	6%	6%	1%	$10,150	$79	$75
	2007	5%	6%	5%	2%	$10,647	$48	$73
	2008	5%	6%	5%	5%	$11,853	$66	$80
	2009	4%	5%	5%	10%	$12,624	$42	$83
	2010	5%	6%	6%	18%	$13,029	$49	$98
	2011	5%	5%	6%	28%	$13,634	$47	$113
	2012	4%	5%	6%	39%	$13,853	$38	$112
	2013	4%	4%	6%	50%	$5,680	$45	$115
	2014	3%	3%	6%	66%	$4,427	$36	$112
	2015	2%	2%	6%	82%	$4,304	$28	$109
	2016	1%	1%	6%	96%	$4,424	$23	$111
All	2004	33%	32%	32%	1%	$19,077	$5,075	$2,958
	2005	33%	31%	31%	1%	$19,463	$4,760	$2,979
	2006	28%	26%	26%	1%	$21,468	$4,893	$2,583
	2007	26%	28%	27%	6%	$22,045	$5,641	$2,823
	2008	25%	26%	26%	13%	$24,578	$5,790	$3,096
	2009	24%	25%	26%	21%	$24,828	$6,363	$3,406
	2010	23%	23%	26%	29%	$26,283	$7,199	$3,690
	2011	21%	21%	25%	37%	$25,915	$5,766	$3,259
	2012	18%	19%	25%	46%	$26,311	$5,286	$2,997
	2013	18%	17%	24%	56%	$17,237	$4,011	$2,763
	2014	16%	16%	25%	64%	$15,530	$3,080	$2,718
	2015	14%	13%	25%	74%	$14,922	$2,220	$2,511
	2016	9%	8%	25%	88%	$14,788	$1,476	$2,577

3

The Role of Incentives in Individual Compliance

Stacy Orlett ◆ Rizwan Javaid ◆ Vicki Koranda
Maryamm Muzikir ◆ Alex Turk

Adam Looney

Maggie R. Jones

Impact of Filing Reminder Outreach on Voluntary Filing Compliance for Taxpayers with a Prior Filing Delinquency

Stacy Orlett, with Rizwan Javaid, Vicki Koranda, Maryamm Muzikir (IRS, Small Business/Self-Employed Division), and Alex Turk (IRS, Research, Applied Analytics, and Statistics)[1]

Introduction

Alternative outreach methods can improve the efficiency of tax administration if the outreach can promote voluntary compliance at a lower cost. Outreach focused on taxpayers at risk of further noncompliance may be able to improve compliance and help the taxpayer avoid unnecessary noncompliance costs. This is particularly important when the Internal Revenue Service (IRS) does not have the resources to intervene with enforcement treatments. The IRS is exploring alternative approaches for promoting voluntary filing, reporting, and paying of tax liabilities. Encouraging the taxpayer to voluntarily meet their tax responsibilities reduces the burden on support and enforcement resources.

In this paper, we report results from IRS outreach that reminded individual taxpayers of their filing obligation. Two separate field tests were conducted exploring alternative methods of promoting voluntary filing compliance among taxpayers with prior filing delinquencies. The alternative treatment included various filing reminder correspondence. Taxpayers treated in the test were sent at least one reminder letter or postcard during the filing season encouraging them to file their return. The first pilot addressed taxpayers where the Automated Substitute for Return (ASFR) process recently brought the taxpayer back into filing by securing a delinquent return. The second pilot addressed taxpayers who had recently not filed a return, and the IRS did not have the resources to take enforcement actions. We found positive effects associated with both the reminder letters and postcards.

Background

The IRS (the Collection organization, working with Research, Applied Analytics, and Statistics (RAAS)) developed two pilot studies to test alternative treatments to promote voluntary filing compliance amongst individuals. Both pilots tested filing reminder correspondence for taxpayers who have had past filing compliance issues.

- **Pilot 1** included taxpayers who went through the ASFR process and whose case was resolved by the taxpayer filing a delinquent tax return in Calendar Year 2015. The objective of this pilot was to determine if the letter and/or postcard could enhance the impact of the ASFR process on subsequent voluntary filing compliance.
- **Pilot 2** included taxpayers who had a requirement to file a Tax Year (TY) 2013 return, but were identified as not having done so in the Individual Case Creation Nonfiler Identification Process (CCNIP), given that the IRS had not started the delinquent return notice process during 2015 due to limited resources.

[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.

Filing compliance issues are addressed in a variety of ways at the IRS. In general, the universe of individual nonfilers will include those who have not filed a tax return by the Return Due Date.[2] The IRS can identify a portion of those with a potential filing requirement after gathering income and other information reported by third parties to the IRS on information returns (e.g., Form W-2).[3] Using this information, the IRS can prioritize and select a portion of these known nonfilers for enforcement action.

The nonfilers who are thus selected, will enter the Return Delinquency notice process and receive up to two notices requesting them to file their tax return. Nonfilers who do not respond may enter the Taxpayer Delinquency Investigation (TDI) process, where they may receive enforcement action from various functions such as the Automated Collection System or call site (ACS), Field Collection (FC), or the ASFR program, to name a few.

Automated Substitute for Return (ASFR)

The ASFR program is highlighted within this paper as it relates to Pilot 1. Cases must meet certain criteria to be assigned to the ASFR program.[4] The ASFR program prioritizes cases by Refund Hold, Tax Year and Net Tax Due.[5] A Refund Hold case has highest priority because the service is holding the taxpayer's refund associated with a recently filed return. The Service has only a limited amount of time to hold the refund. The taxpayer is notified that the Service is holding their refund, and the taxpayer is requested to resolve the delinquent return(s) before the IRS will release the refund. Holding the refund provides a motivation for the taxpayer to file that is different from a non-Refund Hold case. Therefore, Pilot 1 was separated into two groups: (a) Refund Hold ASFR starts; and (b) non-Refund Hold ASFR starts.

If the taxpayer does not respond to the requests by the Refund Hold program to file the delinquent return(s), then the case may be assigned to the ASFR program for additional enforcement. Cases started by the ASFR program are sent a 30-day letter (Letter 2566 or Letter 2566R for non-Refund Hold and Refund Hold cases, respectively). If the taxpayer does not respond timely to the 30-day letter, then a Statutory Notice of Deficiency (also referred to as the 90-day letter) is issued. If the taxpayer does not respond and resolve the issues, then the ASFR program will make a "substitute for return" assessment (default assessment) of tax based on the information documents received from third parties.

Optimally, the taxpayer will respond to the ASFR letters by filing the return or otherwise resolving the issue. This pilot focused on the taxpayers where the ASFR program was successful at securing a delinquent return. The reminder letter or postcard was designed to encourage subsequent voluntary filing compliance.

During Calendar Year 2015, prior to the TY 2015 filing season, we identified approximately 80,000 taxpayers where the ASFR program successfully secured a delinquent return.

- 67 percent were a non-Refund Hold ASFR start,
- 33 percent were a Refund hold ASFR start.

Case Creation Nonfiler Identification Process (CCNIP)

The Case Creation Nonfiler Identification Process[6] is highlighted within this paper as it relates to Pilot 2. The CCNIP is the main process for identifying individual nonfilers with a likely tax filing requirement based on significant income reported by third parties to the IRS, or taxpayers who filed in one year but not in the next (i.e., stop-filers). Due to resource constraints, the IRS is unable to pursue every known case of nonfiling. As mentioned earlier, the pool of nonfilers identified via CCNIP is prioritized using a variety of characteristics or policies, and a portion are selected for enforcement action. Each potential delinquent return is categorized and prioritized based on the case characteristics.[7]

---

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

[3] Internal Revenue Service. Internal Revenue Manual 5.19.2.4.1 (01-16-2015) "IRP Income."

[4] Internal Revenue Service. Internal Revenue Manual 5.18.1.3.1 (12-09-2014) "ASFR Criteria."

[5] Internal Revenue Service. Internal Revenue Manual 5.18.1.3.2 (06-20-2012) "ASFR Prioritization."

[6] Internal Revenue Service. Internal Revenue Manual 5.19.2.2 (01-16-2015) "IMF Return Delinquency Case Creation."

[7] Internal Revenue Service, Document 6209. Section 11–Collection. https://www.irs.gov/pub/irs-utl/6209_section11_2015.pdf (accessed May 2017).

The cases selected can be separated into two general groups based on a Primary Code "B" designation, referred to as PCB and non-PCB in this paper. A PCB case will receive up to two notices and will remain in the notice status if the taxpayer does not respond. The PCB cases generally have a lower priority or risk compared to non-PCB cases. Therefore, Pilot 2 was separated into two groups: (a) PCB; and (b) non-PCB.

After planning and prioritizing the pool of nonfilers and identifying which cases to treat in a given year, the IRS is not always able to start the delinquent return process because of varying workload demands and resource availability. This was true for the Tax Year 2013 cases. Thus, the Return Delinquency program was not able to start as many cases as originally planned. There were approximately 1.9 million nonfilers for the 2013 Tax Year that were identified to be treated but were not, due to changes in resources.

- 16 percent were designated as PCB, and
- 84 percent were designated as non-PCB.

This pool of nonfilers was the focus of Pilot 2. Given that these taxpayers had a potential filing delinquency with their 2013 Tax Year return, this pilot was aimed at identifying a less resource-intensive treatment protocol, which focused on nudging the taxpayer to return to filing timely. The filing reminder outreach was designed to encourage subsequent voluntary filing compliance. The results for Pilot 2 were used to determine if the reminder correspondence could improve voluntary filing compliance on cases where the IRS does not have the resources to take enforcement actions on past noncompliance.

Pilot 1 measured which type of correspondence—letter or postcard—had a bigger impact. Pilot 2 measured how both the message and the frequency of correspondence would impact taxpayer behavior. Pilot 2 had a group that received one postcard and a group that received two postcards.

In addition to these tests, Pilot 2 had two versions of postcards; one postcard was the same one that was used in Pilot 1, while the other postcard had a URL for submitting Form 4506T, which could be used to request tax document transcripts from prior years. The impact of the two different types of postcards was also measured.

Related Research

There is a variety of research from around the world regarding an individual's tax filing compliance. Recent studies explored the impact of preemptive treatments on voluntary filing compliance behavior.

Guyton, *et al.* (2016) was the foundation for our pilots. This study took place during filing seasons in 2014 and 2015, and was comprised of taxpayers who were potentially eligible to claim the Earned Income Tax Credit (EITC) or were otherwise lower-income nonfilers for the 2011 and 2012 Tax Years. There was a control group that received no correspondence and six treatment groups. The treatment groups received correspondence in the form of a preemptive, informative postcard or brochure on voluntarily filing current and prior returns. The correspondence included information on how to file and how to claim the EITC. There were positive results from the treatments with increases to filing rates for those receiving EITC refunds and those who had a balance due. The study also revealed how continued reminders increase the rate of compliance, but that receiving a reminder one year and not the next can potentially lead to recidivism.

Meiselman (2016) conducted a controlled experiment to isolate the effect of message content in mailed communication on income tax nonfiler behavior. The experiment was designed and conducted by the author in collaboration with the City of Detroit. The main outcome of interest was response rate, the rate at which mailings elicited a tax return from taxpayers who had previously failed to file a Tax Year 2014 city income tax return. Other outcomes measured were the amount of remittances, the likelihood of claiming a refund rather than admitting tax due and the number of back-year returns that were filed with the Tax Year 2014 return. The message types included penalty salience, punishment probability, compliance cost, civic pride and the combined penalty salience/punishment probability. The control group consisted of two groups, one group received no contact, and the other group received treatments (postcard and letter) with the prominent message box omitted. This was done to address the phenomenon observed in other studies where taxpayers respond to any contact from the tax agency because they are aware authorities are monitoring their activity.

The author found that "penalty salience" mailings were the most effective at raising response rates, increasing back-year returns filed and returns that admitted tax due. "Punishment probability" was the second most effective. The combined message (punishment probability/penalty salience) was slightly less effective than the penalty salience message alone. The author presented this as evidence supporting the conclusions from prior literature that simplicity is important when communicating with taxpayers. The remaining message types had smaller positive effects, except "civic pride," which had no effect on response rates relative to the basic mailing, which contained no message. The author also did not find evidence of an effect of treatment mailings on the filing behavior of geographic neighbors. This finding was consistent with the notion that networks such as families and coworkers are more relevant to the treated than geographic proximity.

The author also found that mailings were most effective for first-time nonfilers, higher-income taxpayers, and older taxpayers. The author conducted a cost-benefit analysis and determined that there would be a positive improvement to net revenue and welfare if the most effective message treatment—penalty salience—were applied to the entire Detroit income tax nonfiler population. Additionally, net revenue and net welfare could be further improved if the treatment was targeted to higher-yielding demographics (older, higher income, and non-chronic nonfilers).

Kettle (2015) conducted a similar study that took place in Guatemala in early 2014 using preemptive treatments to encourage filing compliance with the 2013 Tax Year. Since tax evasion is a pervasive issue in the country (involving nearly 64% of individuals and businesses), a study with reminder notices was developed to increase voluntary compliance and payments. A control group was created with taxpayers who would not receive any notice, another group received a simple notice to declare/file, and another four groups received letters with different types of behavioral cues. While two of the behavioral letters were the most effective, the simple reminder group also experienced a 3.6 percentage point increase in compliance for an overall 31 percent increase.

A study conducted in the United Kingdom by the Behavioural Insights Team focused on a reminder letter's impact on taxpayers' payment compliance and that they needed to self-assess their tax by a defined date. Taxpayers were separated into three groups based on their past payment compliance: (1) taxpayers who missed their first payment; (2) taxpayers who had been late on payments several times in the past; and (3) taxpayers late on payments in the past, but made their last payment on time. Taxpayers were separated into control and treatment groups. The treatment groups received a letter 10 days prior to the deadline. Groups that received letters all had significantly higher rates of payment. Despite this study focusing on payment compliance, it still drives home the point that there are benefits from receiving a reminder.

The final study we want to highlight was conducted by Chirico (2015) in 2014. The study was conducted in Philadelphia focusing on property owners who were late in paying property tax. Property tax payments are due on March 31st of every year, and if not paid, taxpayers receive a common reminder letter every two months until paid. In this experiment, taxpayers who still had not made any payment by November of that year were selected to receive an alternate reminder. There were four treatment groups: (1) a control group that received the same reminder letter with the addition of a Spanish version; (2) a "Threat" letter that notified taxpayers that their property could be seized or sold by the city; (3) a "Public Service Appeal" letter that notified taxpayers that their tax would be used to fund public services; and (4) a "Civic Duty" letter that notified taxpayers that 9 out of 10 of their neighbors have paid their taxes. While this study focused on payment compliance, it also showed that reminder letters can be effective. Results showed that the "Public Service Appeal" letter was most effective by bringing in an additional $152/letter. The "Civic Duty" letter added $82/letter and the "Threat" letter added $41/letter.

Methodology

We used randomized control trials to test the impact of various preemptive correspondence treatments on filing behaviors of specific ASFR and nonfiler taxpayer groups. Pilot 1 tested a letter versus a postcard and Pilot 2 tested postcard messages and frequency.

Treatments

The following are descriptions of the preemptive correspondence types tested in the experiments:

Letter: The Letter was a generic letter reminding the taxpayer to file a 2015 tax return. The letter included a URL for the IRS website and a toll-free customer service number for the IRS if the taxpayer needed to seek assistance.[8]

Postcard 1 (no Form 4506T reference): Postcard 1 was a generic postcard reminding the taxpayer to file a 2015 tax return. The postcard included a URL for the IRS website if the taxpayer needed to seek assistance.[9]

Postcard 2 (Form 4506T reference): Postcard 2 was a generic postcard reminding the taxpayer to file a 2015 tax return. The postcard included a URL for the IRS website if the taxpayer needed to seek assistance. It differed from Postcard 1 because, in addition to the content of Postcard 1, it also included a URL for the taxpayer where they could submit form 4506T[10] to request tax documents for prior years.[11]

Experimental Design

Pilot 1 tested whether sending a preemptive Letter or Postcard 1 affected the filing behavior of taxpayers previously treated by the ASFR program. Separate experiments were conducted for the Refund Hold and Non-Refund Hold ASFR groups. Random samples were selected to receive a treatment, Postcard 1 or Letter, and a control group sample was selected that was untreated. The correspondence treatments were mailed via the U.S. Postal Service on March 1, 2016.

Pilot 2 tested whether sending preemptive postcards containing different content and the frequency of receiving these postcards affected the filing behaviors of PCB and Non-PCB nonfilers. Separate experiments were conducted for the PCB and Non-PCB nonfiler groups. Pilot 2 treatments also began mailing on March 1, 2016. Treatment groups receiving two postcards were sent a second mailing on April 1, 2016. Like Pilot 1, each group within Pilot 2 also had a control group that was untreated. Random samples of taxpayers were selected to receive one of four types of treatments:

- one mailing of Postcard 1,
- two mailings of Postcard 1,
- one mailing of Postcard 2, and
- two mailings of Postcard 2.

Table 1 lists the treatment and control samples drawn for Pilot 1 and Pilot 2.

TABLE 1. Sample Design for Preemptive Correspondence Pilots

Taxpayer Groups	Pilot 1: ASFR Treated Taxpayers		Pilot 2: TY 2013 Nonfilers	
	Refund Hold	Non-Refund Hold	PCB	Non-PCB
Control	8,142	7,946	7,041	6,550
Letter	8,142	7,946		
Postcard 1 (one mailing)	8,142	7,946	7,041	6,549
Postcard 1 (two mailings)			7,041	6,549
Postcard 2 (one mailing)			7,041	6,549
Postcard 2 (two mailings)			7,041	6,549
Total Sample Size	24,426	23,838	35,205	32,746

[8] An example of the letter can be found in the Appendix.

[9] An example of Postcard 1 can be found in the Appendix.

[10] Form 4506T is a form for requesting tax return information. https://www.irs.gov/pub/irs-pdf/f4506t.pdf (accessed May 2017).

[11] An example of Postcard 2 can be found in the Appendix.

Analysis and Modeling

We conducted regression analysis to estimate the treatment effects of the preemptive contacts on taxpayers timely filing their Tax Year 2015 income tax return or filing for an extension. Data were compiled for each taxpayer in the pilot using the IRS Compliance Data Warehouse (CDW).

The dependent variable for all the models was a dichotomous outcome. A taxpayer received a "1" if the taxpayer timely filed a return for Tax Year 2015 or filed for an extension to file an individual income tax return, and a "0" if not. A "timely filed return" is one that was filed by the return's due date (e.g. April 15, 2016). In addition, we also explored whether the taxpayer filed a prior tax return for years 2011, 2012, 2013 or 2014 during the 2016 filing season.

The independent variables in the models included:

- a treatment dummy for each treatment, respective to the pilot;
- an indicator for undeliverable mail to identify taxpayers with correspondence that came back as undeliverable; and
- a variety of case characteristics derived from information from the Individual CCNIP, the most recent filing information (prior to TY 2015), accounts receivable activity at the beginning of the filing season and one year prior, and delinquent return activity information during the year prior to the start of the filing season.

The treatment dummy variables were defined as "1" if the taxpayer was in the treatment group and had a deliverable mailing address, and "0" if not. The following is a list of the preemptive treatment dummies tested in each pilot:

Pilot 1 (ASFR): 2 Treatment Dummies

1. Postcard 1
2. Letter

Pilot 2 (TY 2013 Nonfilers): 4 Treatment Dummies

1. Postcard 1 (one mailing)
2. Postcard 1 (two mailings)
3. Postcard 2 (one mailing)
4. Postcard 2 (two mailings)

Two-Step Modeling Approach: Accounting for Undeliverable Mail

As with any treatment that involves mailing correspondence to taxpayers, there was a risk that the address on file with the IRS was not accurate, resulting in instances where mail was returned to the Service as undeliverable. In these pilots, there was a portion of taxpayers in each test group whose treatment correspondence was returned to the IRS as undeliverable mail. Unfortunately, we cannot determine which taxpayers in the control group would also have had an invalid address (because they were not sent a mailing). Since invalid addresses do not occur at random and are likely correlated with filing propensity, this can potentially bias our estimated treatment effects.

For both pilots, there were several taxpayers in the treatment groups who had postcards/letters that were returned marked "undeliverable" by the Postal Service (see Table 2). Given that the treatment group was selected at random, we can expect that there are "undeliverable" taxpayers in the control groups, as well, and the taxpayers in the treatment groups would be good surrogates for the "undeliverable" taxpayers in the control.

TABLE 2. Number of Taxpayers with Undeliverable Treatments

Taxpayer Groups	Pilot 1: ASFR		Pilot 2: TY 2013 Nonfilers	
	Refund Hold	Non-Refund Hold	PCB	Non-PCB
Control	n/a	n/a	n/a	n/a
Letter	390	435		
Postcard 1 (one mailing)	450	394	637	576
Postcard 1 (two mailings)			612	508
Postcard 2 (one mailing)			645	570
Postcard 2 (two mailings)			679	553
Total	840	829	2,573	2,207

Our analysis needed to capture the effect of undeliverable mail in the estimated effects of the preemptive correspondence treatments on taxpayer behavior. We used a two-step modeling approach and an instrumental variable for taxpayers with undeliverable mailing addresses. These controls allowed us to estimate unbiased marginal impacts of the preemptive treatments.

For each pilot, we estimated a logistic regression of the likelihood of a taxpayer's correspondence being returned as "undeliverable" among the taxpayers identified for treatment. We used geography as instruments. These model parameter estimates were then used to calculate the "probability of being undeliverable" for all cases (including the control group), and this was included as an independent variable in the second-step logistic regression model, which measured the impact that the preemptive treatments have on taxpayers timely filing TY 2015 returns or extensions.

For the second-step model, a taxpayer with an "undeliverable" preemptive treatment was moved to the control group. This allowed us to compute both the impact of the treatment as well as the marginal impact based on the "intent to treat" the taxpayer.

Findings

The findings are presented for each pilot separately. Because we classified taxpayers as untreated when the correspondence was not delivered, our parameter estimates reflect the impact of the treatment. We also calculated the marginal impact based on the "intent to treat" the taxpayer. This is relevant for tax administration because it reflects the success of the treatment overall.

Pilot 1: Refund Hold and Non-Refund Hold ASFR Taxpayers

Pilot 1 measured the impact of two types of correspondence—a letter and a postcard—on taxpayer behavior among those who filed a delinquent return during the 2015 calendar year following the ASFR process. We found a higher rate of timely filed returns or extensions for the TY 2015 return among the ASFR Refund Hold group compared to the non-Refund Hold taxpayers. Table 3 provides the percentage of taxpayers within each group that timely filed their TY 2015 return or filed for an extension. For the control group, we also computed an adjusted filing rate. This is an estimate of the filing rate for the cases that would have likely received the treatment in the control group by using the undeliverable cases in the treatment group to estimate the number of undeliverable cases in the control group. The adjusted filing rate is calculated as:

$$\text{Adjusted filing rate} = R\,N/N_d - R_u N_u/N_d,$$

where R is the overall rate in the control group, N is the number of cases, R_u is the average rate for the undeliverable cases in the treatment group, N_u and N_d are the estimated number of undeliverable cases in the control and the treatment group, respectively.

TABLE 3. Pilot 1: ASFR—Percentage of Taxpayers with a Timely Filed TY 2015 Return or Extension To File

Taxpayer Groups	Refund Hold		Non-Refund Hold	
	Number of Taxpayers	Percent Timely Filed TY 2015 or Extension	Number of Taxpayers	Percent Timely Filed TY 2015 or Extension
No Treatment	8,982	77.7%	8,775	58.6%
Control	8,142	78.4%	7,946	59.9%
Control (adjusted filing rate)		78.8%		60.7%
Undeliverable	840	71.1%	829	46.0%
Letter	7,752	81.4%	7,511	64.2%
Postcard 1 (one mailing)	7,692	79.6%	7,552	61.0%

SOURCE: IRS. Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

Table 4 contains the regression results for ASFR Refund Hold and non-Refund Hold taxpayers in our study. In this regression, we measured the effect of receiving a post card or a letter on the timely filing a Tax Year 2015 return or filing for an extension.

For both the Refund Hold and non-Refund Hold groups, there were significant, positive impacts of both the postcard and the letter on voluntary filing. There was just slightly more than a one percentage point increase in the propensity to file when a taxpayer received a postcard. However, receiving a reminder letter to file resulted in a statistically significant increase in the propensity to file the TY 2015 return that was twice as large (relative to a postcard) for the Refund Hold cases and three times as large for the Non-Refund Hold.

There is anecdotal evidence that many taxpayers with compliance issues don't open correspondence from the IRS. Thus, sending a postcard could overcome the "fail to open" barrier when nudging taxpayers towards compliance. These results are not consistent with that notion. One possible explanation is that a barrier exists for taxpayers who know they have an issue. Since these ASFR taxpayers in the study came back into filing compliance, they may have been alarmed when they received a letter because they did not think they had an issue. Alternatively, we have no way of knowing if the taxpayer opened the letter. It could be that receiving the letter prompted the taxpayer to respond, and the postcard was more likely to go unnoticed.

TABLE 4. Pilot 1 Regression Results for ASFR Taxpayers Dependent Variable: Taxpayer Timely Filed TY 2015 or Filed for an Extension

Explanatory Variables	Refund Hold			Non-Refund Hold		
	Estimate	Marginal Effect of Treatment	Marginal Effect of Intent to Treat	Estimate	Marginal Effect of Treatment	Marginal Effect of Intent to Treat
Intercept	-0.354*			-0.515*		
	(0.116)			(0.066)		
Post Card Treatment	0.099*	0.013	0.013	0.083*	0.011	0.010
	(0.040)			(0.035)		
Letter Treatment	0.198*	0.027	0.025	0.244*	0.033	0.031
	(0.041)			(0.035)		
Probability Undeliverable Mail	-2.519*			-4.878*		
	(1.088)			(0.754)		
Log Likelihood	-11,435			-13,972		
Number of observations	24,368			23,431		

* Indicates significance at the 95% level.
NOTE: Not all explanatory variables are shown. Standard errors are reported in parentheses.
SOURCE: IRS, Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

Pilot 2: Nonfiler PCB and Non-PCB Taxpayers

Pilot 2 was conducted to explore how the content and the frequency of the preemptive correspondence can improve voluntary filing compliance on cases that were potential nonfilers in the past but have not been contacted. The analysis is conducted separately for PCB and non-PCB potential nonfilers. There were four treatments:

- Treatment A: Postcard 1 (one mailing, *same version used for Pilot 1*)
- Treatment B: Postcard 1 (two mailings, *same version used for Pilot 1*)
- Treatment C: Postcard 2 (one mailing, *version with Form 4506T reference*)
- Treatment D: Postcard 2 (two mailings, *version with Form 4506T reference*)

We found a similar rate of timely filed returns or extensions for TY 2015 between the PCB and the non-PCB TY 2013 potential nonfilers. Table 5 provides the percentage of taxpayers within each group who timely filed their Tax Year 2015 return or filed for an extension.

TABLE 5. Pilot 2: Percentage of Taxpayers with a Timely Filed TY 2015 Return or Extension to File Among TY 2013 Potential Nonfilers

Taxpayer Groups	Primary Code B		Non-Primary Code B	
	Number of Taxpayers	Percent Timely Filed TY 2015 or Extension	Number of Taxpayers	Percent Timely Filed TY 2015 or Extension
No Treatment	9,614	42.6%	8,757	45.9%
Control	7,041	46.5%	6,550	49.7%
Control (adjusted filing rate)		47.9%		51.1%
Undeliverable	2,573	32.2%	2,207	34.9%
Postcard 1 (one mailing)	6,404	49.0%	5,973	52.0%
Postcard 1 (two mailings)	6,429	49.3%	6,041	53.0%
Postcard 2 (one mailing)	6,396	49.2%	5,979	51.9%
Postcard 2 (two mailings)	6,362	49.5%	5,996	51.9%

SOURCE: IRS, Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

Table 6 shows the regression results for the potential Tax Year 2013 nonfilers receiving treatments A, B, C and D. For **PCB taxpayers**, all the estimates were positive, but Treatment A was not statistically significant; however, Treatment B (sending two of the Pilot 1 postcards) was statistically significant. Both mailing frequencies for Postcard 2 positively impacted filing propensity, but again sending two had a larger impact. This suggests that multiple contacts may be needed for the lower priority cases.

For **non-PCB taxpayers**, all treatments resulted in a statistically significant increase in the number of delinquent returns filed by these taxpayers. However, it appears there was less of an impact with a second postcard. Interestingly, the Pilot 1 postcard appears to be more effective. The Postcard 2 contained more information and referenced how to request transcripts of prior tax years. Thus, the nudge toward filing appears to be more effective without the extra information. It could be that the Postcard 1 message was clearer to the taxpayer and thus more effective. It also could be that alluding to past potential noncompliance makes the taxpayer more hesitant to file, perhaps because they are afraid they will then be contacted about the unfiled return. This is a behavioral response that could be tested further.

TABLE 6. Pilot 2 Regression Results for Tax Year 2013 Potential Nonfilers Dependent Variable: Taxpayer Timely Filed TY 2015 or Filed for an Extension

Explanatory Variables	Primary Code B			Non-Primary Code B		
	Estimate	Marginal Effect of Treatment	Marginal Effect of Intent To Treat	Estimate	Marginal Effect of Treatment	Marginal Effect of Intent To Treat
Intercept	-1.393*			-1.916*		
	(0.055)			(0.139)		
Treatment A—Postcard 1	0.0589	0.010	0.009	0.110*	0.017	0.016
(1 mailing, same version used for Pilot 1)	(0.041)			(0.042)		
Treatment B—Postcard 1	0.1447*	0.024	0.021	0.140*	0.022	0.020
(2 mailings, same version used for Pilot 1)	(0.041)			(0.041)		
Treatment C—Postcard 2	0.1038*	0.017	0.015	0.087*	0.014	0.013
(1 mailing, version refers to Form 4506T)	(0.041)			(0.042)		
Treatment D—Postcard 2	0.1293*	0.021	0.019	0.084*	0.013	0.012
(2 mailings, version refers to Form 4506T)	(0.041)			(0.042)		
Probability of Undeliverable Mail	1.878*			-2.086*		
	(0.422)			(0.494)		
Log Likelihood	-16,629			-16,320		
Number of observations	33,259			32,421		

* Indicates significance at the 95% level.
NOTE: Not all explanatory variables are shown. Standard errors are reported in parentheses.
SOURCE: IRS, Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

Table 6 also shows that PCB taxpayers for whom we don't have a good address (i.e., *Probability of Undelivered Mail*) had a statistically significant positive impact on whether the taxpayer filed a delinquent return in TY 2015. This is counter-intuitive. This may be because these taxpayers have already been predetermined to not be a high-priority case (hence the PCB designation) because they were less likely to have a filing requirement. It seems possible that there are more bad addresses for cases where the taxpayer does not have a requirement to file.

Prior Tax Return Filings

An ancillary benefit of the preemptive outreach is that some taxpayers may also file past delinquent returns. Thus, we identified prior tax returns that were filed during the TY 2015 filing season. A prior return was identified using data on the IRS Compliance Data Warehouse through the end of August 2016 (or cycle 34). Any taxpayer from a treatment group marked as undeliverable was moved to the Control group for this exercise. This helped identify the impact of the treatment on filing prior tax returns.

By the nature of the ASFR Refund Hold group, most of these taxpayers had recently filed a return for which the IRS held a refund because the taxpayer had an unfiled return for a different year. These taxpayers may likely not have as many prior returns to file compared to the other groups. In the non-Refund Hold ASFR group, we found a significant difference between the control and postcard groups in filing a prior return.

TABLE 7. Pilot 1: ASFR—Percentage of Taxpayers with at Least One Prior (TY 2011–2014) Tax Return During the TY 2015 Filing Season

Taxpayer Groups	Refund Hold		Non-Refund Hold	
	Number of Taxpayers	Percent Filed at Least One TY 2011–TY 2014	Number of Taxpayers	Percent Filed at Least One TY 2011–TY 2014
No Treatment	8,982	4.0%	8,775	11.6%
Control	8,142	3.8%	7,946	11.4%
Control (adjusted filing rate)		3.7%		11.3%
Undeliverable	840	5.5%	829	13.3%
Letter	7,752	3.8%	7,511	12.0%
Postcard 1 (one mailing)	7,692	3.8%	7,552	12.7%

SOURCE: IRS, Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

A significant difference exists across all treatments in Pilot 2 involving the TY 2013 potential nonfilers, with a slightly larger percentage point difference among the PCB group. Thus, the simple comparison suggests that outreach may also be effective in resolving past delinquencies.

TABLE 8. Pilot 2: Percentage of Taxpayers Among TY 2013 Potential Nonfilers with at Least One Prior (TY 2011–2014) Tax Return During the TY 2015 Filing Season

Taxpayer Groups	Primary Code B		Non-Primary Code B	
	Number of Taxpayers	Percent Filed at Least One TY 2011–TY 2014	Number of Taxpayers	Percent Filed at Least One TY 2011–TY 2014
No Treatment	9,614	11.5%	8,757	14.6%
Control	7,041	12.1%	6,550	15.8%
Control (adjusted filing rate)		12.3%		16.2%
Undeliverable	2,573	9.7%	2,207	11.2%
Postcard 1 (one mailing)	6,404	13.7%	5,973	17.1%
Postcard 1 (two mailings)	6,429	13.7%	6,041	16.0%
Postcard 2 (one mailing)	6,396	13.9%	5,979	16.3%
Postcard 2 (two mailings)	6,362	13.6%	5,996	17.0%

SOURCE: IRS, Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

Table 9 shows the percentage point difference between the portions of taxpayers in each treatment group filing at least one prior tax return (Tax Years 2011–2014) during the TY 2015 filing season compared to the untreated group.

TABLE 9. Percentage Point Difference From the No Treatment Group in the Portion of Taxpayers Filing at Least One Prior (TY 2011–2014) Tax Return During the TY 2015 Filing Season

Taxpayer Groups[a]	Pilot 1: ASFR Treated Taxpayers		Pilot 2: TY 2013 Nonfilers	
	Refund Hold	Non-Refund Hold	PCB	Non-PCB
Letter	-0.2	0.4		
Postcard 1 (one mailing)	-0.2	1.1*	2.3*	2.5*
Postcard 1 (two mailings)			2.3*	1.3*
Postcard 2 (one mailing)			2.4*	1.6*
Postcard 2 (two mailings)			2.2*	2.4*

[a] Taxpayers with undeliverable treatments were moved to the Control group.
* Denotes a significant difference from the control group at the 95% level.
SOURCE: IRS, Compliance Data Warehouse. Individual Return Transaction File, Individual Master File Status and Transaction History, Individual Accounts Receivable Dollar Inventory, and Case Creation Nonfiler Identification Process. Data Extracted September 2016.

Conclusions and Direction for Further Research

In this paper, we showed that preemptively contacting taxpayers who have had a previous filing compliance issue can improve future filing compliance. The impacts were modest, but these impacts came at a very low cost. Our results suggest that a letter, at least for some taxpayers, may be more effective than a postcard. When using postcards to nudge taxpayers, lower risk taxpayers may need multiple nudges for the treatment to be effective. Our results also supported the notion that a clearer message may be more effective in increasing taxpayer response, at least in terms of voluntary filing. Also, we plan to extend the analysis of the impact of outreach on past compliance.

There are several areas where these results need to be extended. We need to fully understand the differing results from letters vs. postcards in a broader population of nonfilers. Previous work on EITC suggested that letters vs. postcard was not important, but perhaps it is important for the general population of potential nonfilers. Also, is the "opening" of the letter a real barrier or does just receiving the letter, even if unopened, have an impact on behavior?

The results also need to be extended further to understand to what degree, if any, trying to address past nonfiling is an impediment to fostering compliance in the future. If the tax authority does not have the resources to go back and enforce filing, is it better for them to "cut their losses" and focus only on the taxpayer's future filing behavior? These behavioral tests for filing compliance and compliance in general could be addressed in field and potentially in laboratory experiments to explore the impact of policies that ignore past noncompliant behavior and rather focus on taxpayers making a fresh start toward compliance.

References

Chirico, M., R. Inman, C. Loeffler, J. MacDonald, H. Sieg. (2015). "An Experimental Evaluation of Notification Strategies to Increase Property Tax Compliance: Free-Riding in the City of Brotherly Love." In: J.R. Brown, editor. *Tax Policy and the Economy*. Volume 30. Chicago, United States: National Bureau of Economic Research and University of Chicago Press.

Datta, S., S. Orlett, A. Turk. (2015). "Individual Nonfilers and IRS Generated Tax Assessments: Revenue and Compliance Impacts of IRS Substitute Assessment When Taxpayers Don't File," *2015 IRS Research Bulletin,* Publication 1500, pp. 113–137.

De La Matta, J. C., J. Guyton, R. Hodge II, P. Langetieg, S. Orlett, M. Payne, A. Qadri, L. Rupert, B. Schafer, A. Turk, M. Vigil. (2016). "Understanding the Nonfiler/Late Filer: Preliminary Findings." *2016 IRS Research Bulletin,* Publication 1500, pp. 140–162.

Hallsworth, Michael (2016). "7. Reducing Fraud, Error and Debt." *The Behavioural Insights Team's Update Report: 2015-16.* pp. 55–58.

Greene, William. (2003). *Econometric Analysis.* New Jersey: Prentice Hall.

Guyton, J., D. Manoli, B. Schafer, M. Sebastiani. (2016). "Reminders and Recidivism: Using Administrative Data to Characterize Nonfilers and Conduct EITC Outreach." *NBER Working Paper Series,* Working Paper 21904 (http://www.nber.org/papers/w21904, last accessed 09/21/17).

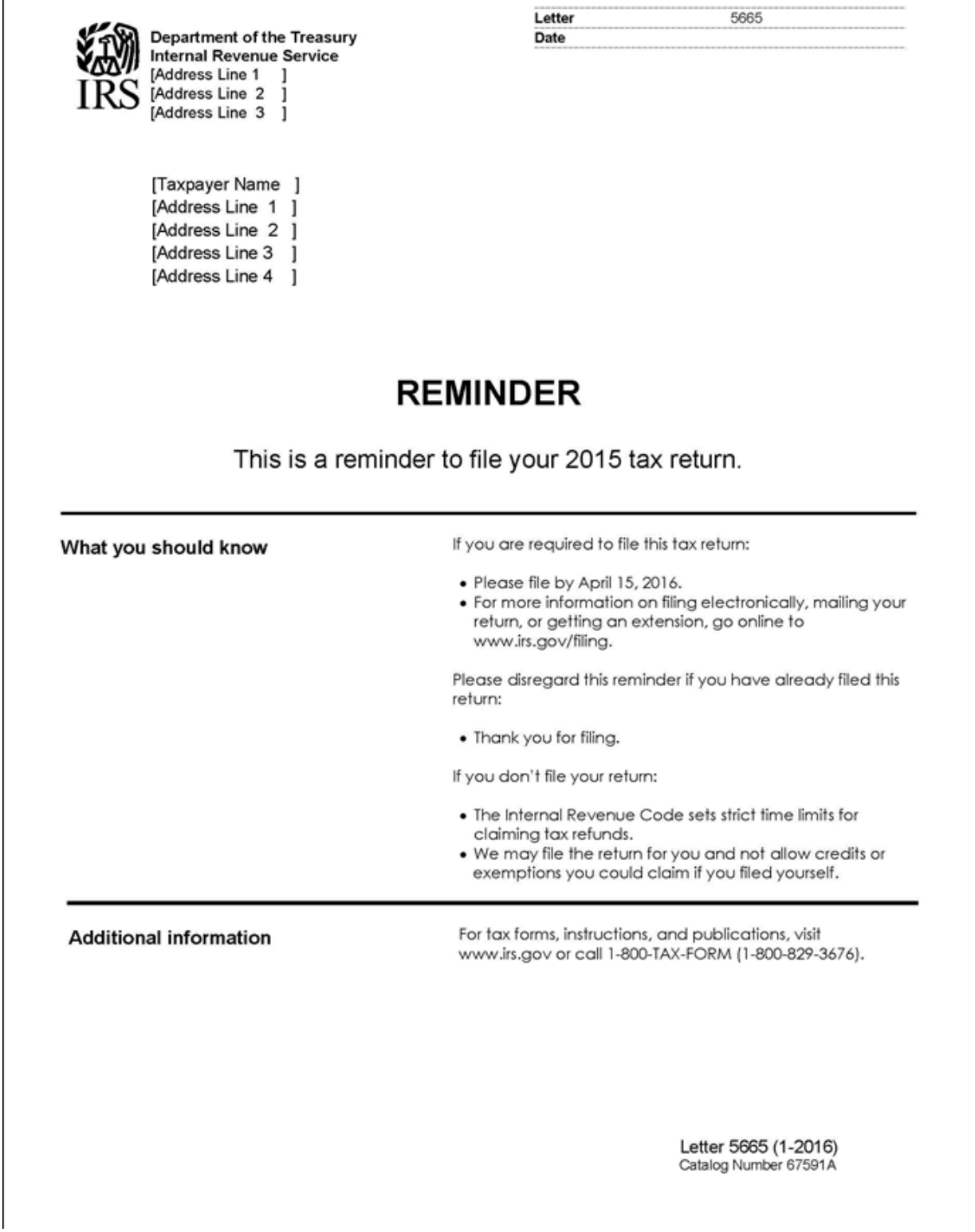
Kettle, S, M. Hernandez, S. Ruda, M. Sanders. (2015). "Behavioural Interventions in Tax Compliance: Evidence from Guatemala." *2015 IRS Research Bulletin*, Publication 1500, pp.141–165.

Meiselman, Ben. (2016). "Ghostbusting in Detroit: Evidence on nonfilers from a controlled field experiment." University of Michigan.

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).

Appendix—Samples of Treatment Correspondence

Preemptive Letter (P5665)



Department of the Treasury
Internal Revenue Service
[Address Line 1]
[Address Line 2]
[Address Line 3]

Letter	5665
Date	

[Taxpayer Name]
[Address Line 1]
[Address Line 2]
[Address Line 3]
[Address Line 4]

REMINDER

This is a reminder to file your 2015 tax return.

What you should know

If you are required to file this tax return:

- Please file by April 15, 2016.
- For more information on filing electronically, mailing your return, or getting an extension, go online to www.irs.gov/filing.

Please disregard this reminder if you have already filed this return:

- Thank you for filing.

If you don't file your return:

- The Internal Revenue Code sets strict time limits for claiming tax refunds.
- We may file the return for you and not allow credits or exemptions you could claim if you filed yourself.

Additional information

For tax forms, instructions, and publications, visit www.irs.gov or call 1-800-TAX-FORM (1-800-829-3676).

Letter 5665 (1-2016)
Catalog Number 67591A

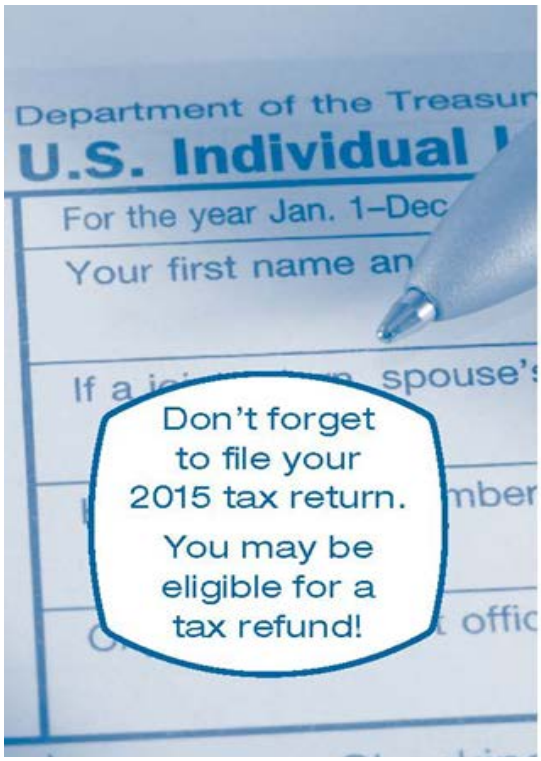
Preemptive Postcard 1(P5235)

(front)

If you have not already done so, remember to file your 2015 tax return by April 15, 2016.

- Did you know the average tax refund in 2014 was approximately $2,800?
- You could be eligible for valuable tax benefits, but you must file to receive them.
- For more information about filing, or getting an extension to file, go online to www.irs.gov/filing.

(back)



IRS

Internal Revenue Service
c/o Westat
1600 Research Blvd, RW2634
Rockville, MD 20850-3129
RETURN SERVICE REQUESTED

Taxpayer's Address Here

Publication 5235 (2-2016) Catalog Number 68313K
Department of the Treasury **Internal Revenue Service**
www.irs.gov

Preemptive Postcard 2 (P5235A)

(front)

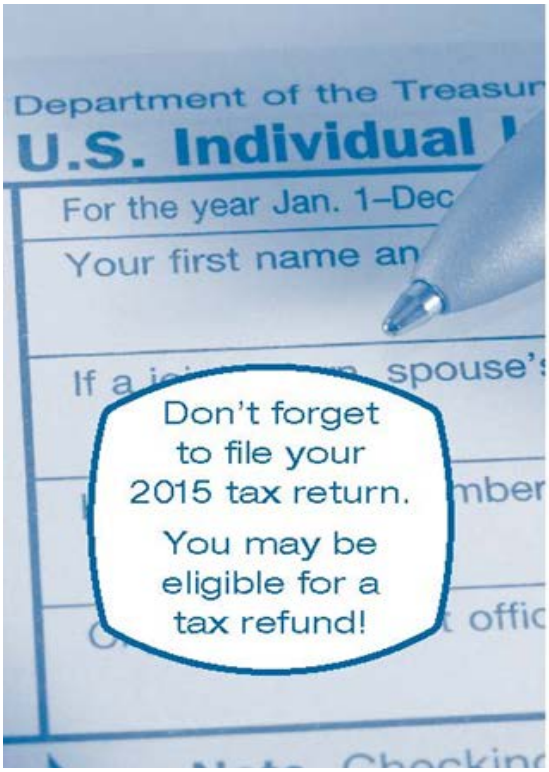
If you have not already done so, remember to file your 2015 tax return by April 15, 2016.

- Did you know the average tax refund in 2014 was approximately **$2,800?**
- You could be eligible for valuable tax benefits, but you must file to receive them.
- For more information about filing, or getting an extension to file, go online to www.irs.gov/filing.

It's not too late to file returns for prior tax years.

- You can file late tax returns and claim tax refunds up to 3 calendar years after the April filing deadline. For example, you can claim a tax refund for 2012, if you file your tax return by April 15, 2016.
- To request transcripts of prior year W-2s and other tax documents, submit Form 4506T (from http://www.irs.gov/uac/About-Form-4506T).

(back)



IRS

Internal Revenue Service
c/o Westat
1600 Research Blvd, RW2634
Rockville, MD 20850-3129

RETURN SERVICE REQUESTED

Taxpayer's Address Here

Publication 5235-A (2-2016) Catalog Number 68538H
Department of the Treasury **Internal Revenue Service**
www.irs.gov

Charitable Contributions of Conservation Easements

Adam Looney (Brookings Institution) [1]

A tax deduction intended to encourage conservation of environmentally important land and historic buildings has become a lucrative way for real estate developers to finance development projects—reducing their tax payments by billions of dollars, and in some cases doing little to advance environmental protection.

Created 40 years ago, the provision allows property owners to take a charitable deduction for donating qualified conservation easements—legal agreements that permanently limit the development or use of a property—to a charitable organization. The deduction has proved to be a popular and successful tool for encouraging the conservation of environmentally important land and historic buildings, and the tax benefit to donors is often seen as a key component in making a conservation deal come together. However, this obscure tax provision has proven difficult to administer and enforce, and ranks among the top 10 most litigated issues between the IRS and taxpayers (National Taxpayer Advocate (2015)). Some donors abuse the provision by applying grossly inflated appraisals to the value of the easement to increase their charitable deduction or by taking donations for easements that do not fulfill bona fide conservation purposes.

Indeed, some real estate developers exploit these vulnerabilities by selling the rights to claim charitable deductions to investors and using the proceeds to finance development, which costs taxpayers hundreds of millions of dollars per year and undermines the program's conservation goals. In these transactions, developers promote arrangements structured to provide investors with a "return" in the form of inflated charitable deductions, sometimes well in excess of the value of their initial investment. The developer will use the initial financing to purchase the land, make improvements or change zoning rules, and develop part of the property (like building condominiums or a club house). The improvements are then used to justify a larger appraisal on an easement on the remaining open space. Because of how some donee organizations report donations (or fail to do so) the magnitude of these abuses is hidden from public scrutiny. But at least three of the five largest donee organizations (by contribution volume) appear to actively participate in these arrangements.

The dollar amounts attributable to such abusive transactions appear to have surged in recent years. Total deductions for conservation easement contributions by taxpayers tripled in 2014—rising from $971 million in 2012 to $1.1 billion in 2013, and to $3.2 billion in 2014, according to preliminary IRS tabulations. Syndications—the selling off of deductions to investors—appear to be one source of this surge. Data for 2015, 2016, and 2017, years in which promoted syndications appear to have become more prevalent, are not yet available.

Concerns about abuses of conservation easements predate this recent surge. Among the concerns:

Donations are concentrated in transactions that seem unrelated to conservation benefits

> The dollar value of donations of conservation easements, though not acreage donated nor the number of properties, is highly concentrated in certain types of transactions, in certain geographic areas, and in a handful of donee organizations. For instance, between 2010 and 2012, about 36 percent of all deductions nationwide for donations of conservation easements were taken by taxpayers in Georgia. According to the Land Trust Alliance, Georgia has 1.5

[1] This paper was prepared, circulated, and submitted for presentation at the IRS-TPC research conference while the author was a Treasury employee. I am grateful to Mark Mazur, Elinor Ramey, and many other colleagues for helpful comments and discussions, and especially to Ruth Madrigal for sharing her passion, wisdom, and good ideas. This research relies on the analytical capability that was made possible in part by a grant from the Laura and John Arnold Foundation. The findings and conclusions contained within are those of the author and do not necessarily reflect positions or policies of the Tax Policy Center or its funders.

percent of conserved land. Connecticut, which is smaller in size than all but two States, ranked third in easement deductions. It had 7 percent of all conservation-easement deductions, but only 0.4 percent of the acres under easement because land in the wealthiest parts of that State is so valuable. (See Table 3.)

About 10 percent of the acreage under easement claims about 69 percent of all tax benefits, largely because the valuation of the easements (per transaction or on a per-acre basis) is unusually high. Similarly, the 2 percent of transactions between 2010 and 2012 valued over $5 million each, account for 43 percent of all deductions. A relatively small number of taxpayers, transactions, and donee organizations reap a large share of the total tax benefit.

The concentration of donations is associated with two factors: (1) easements related to large real estate developments, such as tract housing, private communities, or recreational facilities, like country clubs and golf clubs; and (2) donations in high-cost areas, like expensive suburbs of major metropolitan areas (e.g. Atlanta, GA, or Westchester, NY) or vacation or resort destinations (like Jackson Hole, WY, or the Cape and Islands, MA). Among the roughly two dozen transactions recently examined by the IRS involving easements to conserve open space on a property that included a golf course, the average charitable deduction claimed by the owner was $19 million. Taxpayers justify the large deduction on the basis that the improvements to the land, the proximity to high-cost residential property, and other amenities result in high land values were the land developed for another use.

When private charities and Federal and State elected officials allocate spending to purchase or conserve land, they do not spend the vast majority of their resources to preserve golf courses, suburban subdivisions, real estate developments, or vacation homes. The disparity in where the money goes suggests that the tax expenditure is not flowing to preserve properties with high conservation value.

A small handful of donee organizations are responsible for a disproportionate share of donations

Between 2010 and 2012, some 25 organizations (of about 1,700 land trusts nationwide) received about half of all donations of easements, measured in dollar value. A few of these are large, nationally-recognized organizations whose conservation efforts are transparently documented and communicated in their public filings. Many, however, are small organizations with few employees and scarce management or enforcement resources. And most do not report receiving gifts of easements or do not report the value of the easements they receive in public filings.

For instance, according to publicly available data from the IRS, the largest recipient of donations of easements by dollar value over the period from 2011 to 2013 was the Foothills Land Conservancy of Maryville, TN. The organization reported having 4 employees and spending $19,000 to monitor the 19,600 acres of easements it maintains in 5 States. Foothills received 14 contributions of easements valued at $236.7 million (averaging about $17 million each) in 2013. That is roughly a quarter of the total volume of deductible donations of qualified conservation easements claimed by all individual taxpayers in 2013 ($1.1 billion). Based on the total value of deductible contributions that year, Foothills would rank alongside America's largest public charities, according to Forbes annual ranking, "The Largest U.S. Charities for 2016" (Forbes 2016). (In 2015, Foothills stopped reporting the value of noncash gifts and its reported revenues fell by 99.8 percent.) By contrast, the second largest recipient of donations of easements was the Nature Conservancy, which maintains almost 2.9 million acres, has 3,725 employees, and spends more than a million dollars each year maintaining and enforcing their easements. According to their IRS return for 2011, they received 76 easements valued at $95 million total.

Most organizations that receive donations of easements do not report them as gifts or revenues on their public tax returns

The tax returns of charitable and tax-exempt organizations are public to provide information about the activities of the charitable sector, to provide transparency and accountability, and to help reduce any abuse of tax-exempt status. The amount of gifts (or revenues) that an organization receives is a critical component of the return because it illustrates the scale of the organization, the size of the "tax expenditure" its deductible gifts represent, and the way specific rules are implemented regarding charities' public support.

Many of the organizations that manage hundreds of millions of dollars in conservation-easement donations each year cannot be identified in public records because they do not disclose the value of gifts of easements on their tax returns. Often, they report the value of easement donations at zero. Were these organizations to include these easement donations at appraised value (as they would if they received gifts of cash or marketable securities), several organizations that currently appear to be small organizations would rank among the nation's 100 largest nonhospital, nonuniversity charitable organizations.

In addition to impeding the transparency and public accountability intended by public disclosure of charities' tax returns, this convention may also allow organizations that should be legally classified as private foundations to qualify as public charities by subverting a test related to the breadth of the organizations' public support.

Donations of "partial interests" are difficult to administer

The difficulty in administering this provision—and its vulnerability to abuse—arises in part because of the unusual nature of donations of conservation easements. These donations transfer only certain rights (such as the right to develop the land) to the donee organization, while the donor (landowner) retains ownership and certain other rights to continue to use the property. In most other circumstances, a donor must contribute his or her entire interest in donated property to take a deduction; donations of only some of an owner's property rights are generally not deductible. The tax law allows a deduction for the fair market value of a qualified conservation contribution. Appraising the value of the partial interest (separately from the remaining interest), which is necessary for determining the taxpayer's deduction, has proved contentious and is the source of much of the litigation between taxpayers and the IRS. Because the donee generally cannot monetize the value of the contribution by selling the property, as it could with gifts of stock or other property, there is no guarantee that the value of the deduction claimed by the donor is commensurate with its value to the donee. High-profile examples of taxpayers taking large charitable donations for conservation easements on properties like golf courses have fueled concerns that some donations of easements have benefitted donors more than they have furthered conservation causes (see, e.g., Rubin (2016b)).

What Can Be Done?

Abuse of the conservation easement deduction reduces tax revenue, raises the appearance of unfairness and inequity in the tax system, hinders conservation goals, and causes a disproportionate amount of IRS enforcement and taxpayer burden. Beyond the cases of abuse, a key policy question is whether this tax expenditure represents a good return on the scarce dollars the Federal Government uses to subsidize it. It is clear that privately funded conservation organizations do not want to prioritize or finance the preservation of golf courses or the grassy areas between tract housing when spending their own money. Instead, they buy undeveloped land with special environmental or public recreational values. This provides one indication that the current structure is inefficient—if environmental organizations, rather than taxpayers, could allocate how the money was spent, we would expect a very different pattern of conservation.

Policy changes could reduce the incidence of abuse, reduce (or redirect) the tax expenditure, improve transparency and accountability, improve the conservation value achieved with the tax benefit, or some

combination thereof. With policy fixes, we could get more conservation for the tax expenditure we provide, helping to achieve conservation goals while minimizing the appearance of unfairness and the number of abuses.

Make promoted, syndicated easement transactions a "listed transaction"

A recent IRS notice was an important first step for promoting transparency and in identifying transactions with the largest propensity for abuse (IRS (2017)). The notice requires participants in promoted, syndicated transactions that promise charitable deductions far in excess of their investments to "raise their hand" so that the IRS and policymakers can understand the scope of these transactions and to understand whether these transactions require greater scrutiny. Promoters of these investment deals have since lobbied to have the listing notice rescinded—and have won a delay of the implementation of the listing notice until later in 2017. But the notice does not change the law or regulations related to conservation easements, holds harmless donee organizations, and has no effect on the vast majority of traditional easement donations. The listing notice should remain in force and the IRS should implement it on schedule.

Increase transparency

Publicly available returns of donee organizations frequently exclude information on the value of donations of easements, either because the donee organizations do not report the value of the charitable contributions of easements they receive on their tax returns or value them at zero. For instance, of the 21 largest nongovernmental recipients of charitable donations of conservation easements between 2010 and 2012 identified by examining donor records, 15 organizations did not disclose the value of donations received. This practice makes it impossible to know which organizations are active in using this incentive, impairing the ability of the general public, tax administrators, or associations of land trusts to monitor the activity of this sector. Moreover, it may also allow some organizations to circumvent the public charity support test, which generally requires that an organization receive at least one-third of its support from contributions from the general public, or meet a 10-percent "facts and circumstances" test, or to qualify to file Form 990-EZ or Form 990-PF, which reduces the amount of information disclosed.

Several options to increase reporting by donors (on Form 8283, where they claim donations) and by donee organizations (on Form 990 and its supplemental schedules) would provide "sunshine" to help the public understand.

Strengthen the definition of conservation purpose and standards for organizations

While tax benefits for conservation easements were intended to be available only for certain conservation purposes, the scope of what qualifies as a valid purpose has expanded to include easements on properties that do not provide public benefits or do not further bona fide conservation policies. All donations of easements should both fulfil a clearly delineated conservation policy (or an authorized State or tribal policy) and yield a significant public benefit. In addition, both the donor and donee organization should attest to the public benefit and conservation purpose and provide a justification thereof in the course of claiming any tax benefits.

Receiving and preserving conservation easements in perpetuity (as the law requires) is a costly and burdensome responsibility. Recognizing these challenges, several States and voluntary accreditation programs have developed minimum requirements for organizations to qualify to receive donations of easements. Such qualifying minimum standards should also apply to organizations entrusted with federally subsidized conservation easements.

Some States and Federal agencies appoint boards to preapprove easements before any benefits are provided to the landowner to ensure the easement achieves a stated conservation purpose

and the appraisal is appropriate. For example, Colorado established its Conservation Easement Oversight Commission and modified its tax credit program in the wake of similar abuses.

Use an allocated credit instead of a deduction

A more fundamental reform of tax subsidies for conservation easements would take the deduction and transform it into a credit allocated to donee organizations. In this model, donee organizations would be empowered to approach landowners to "spend" the credit and to decide what kind of properties to conserve and how much to pay. Because these organizations have the right incentives to conserve properties with the greatest environmental or historic value, this approach is intended to maximize the return on the tax benefits provided without requiring adversarial IRS oversight.

History and Background

Tax benefits for gifts of conservation easements were first enacted in the Tax Reform Act of 1976 and the Tax Reduction and Simplification Act of 1977, and extended permanently in 1980 with the Tax Treatment Extension Act, which created IRC section 170(h). The Taxpayer Relief Act of 1997 extended the deductibility of conservation easements against the estate tax. Rules made permanent in the Protecting Americans from Tax Hikes Act of 2015 enhanced the tax benefits available for donations of easements by allowing individuals who make donations of conservation easements to deduct up to 50 percent of their contribution base (generally, AGI) and individuals who are qualified farmers and ranchers may deduct up to 100 percent of their contribution base.

In order to be deductible, a donation of a conservation easement must be made to a qualified charitable organization—generally either a public charity or a State or local government entity (but excluding, for instance, private foundations).[2] The contribution must be exclusively for a conservation purpose, where the term "conservation purpose" means:

- the preservation of land areas for outdoor recreation by, or the education of, the general public;
- the protection of a relatively natural habitat of fish, wildlife, or plants, or similar ecosystem;
- the preservation of an historically important land area or a certified historic structure; or
- the preservation of open space (including farmland and forest land) where such preservation is:
  - for the scenic enjoyment of the general public, or pursuant to a clearly delineated Federal, State, or local governmental conservation policy; and
  - will yield a significant public benefit.

Evidence of the Use and Abuse of Conservation Easements in IRS Data

The analysis of data from a variety of IRS sources provides a picture of how conservation easements are used and sometimes abused. These data show that easements are highly concentrated in a relatively small number of large-dollar transactions, in certain geographic areas, and in a small number of organizations. This pattern of donations corroborates concerns voiced in the press and by advocates over large donations taken for properties, like golf courses, with questionable conservation values, and the red flags raised by IRS enforcement officials and the Taxpayer Advocate, focused on the high rates of litigation over easement deductions.

Data on conservation easements come from two sources: the return of charitable organizations, who are required to report donations of easements in various places on their publicly available return (Form 990), and

[2] To qualify as a public charity, an organization must have broad public support. Most charities qualify by receiving at least one-third of their support from contributions from the general public and/or from receipts from activities related to their tax-exempt purpose. Organizations that fail to meet that threshold—such as when one individual, family, or organization provides a large share of the contributions—are private foundations and governed by stricter rules intended to prevent against self-dealing.

Form 8283, *Noncash Charitable Contributions,* which must be filed by individual taxpayers who claim itemized deductions for donations of easements.

For context, according to the latest IRS report *Individual Noncash Contributions, 2013,* some 2,025 taxpayers reported making contributions of conservation easements on their 2013 returns, some of them with more than one donation. In total, these taxpayers deducted $1,083,785,000 of contributions, an average of about $535,311 per taxpayer (IRS (2016)). These statistics are derived from tabulations of information from Form 8283 (the form taxpayers must submit to document certain noncash donations) from a stratified random sample of individual taxpayers.[3] Preliminary tabulations from 2014 show that 3,249 taxpayers claimed total contributions of easements of $3.2 billion (about $983,729 per return).

Detailed Information from Form 8283

To provide more detailed information on donations of conservation easements, we examined information drawn from Form 8283 for a sample of taxpayers from Tax Years 2010, 2011, and 2012. These data are based on a stratified sample of taxpayers in which high-income taxpayers (those likeliest to contribute deductions for easements) are overrepresented. The sample population generally corresponds to the samples used to inform the IRS SOI reports on noncash contributions referenced above. However, the sample is slightly larger because we expanded the number of transactions to include any donation whose description includes the word "easement," which results in a slightly larger sample of taxpayers, donations, and dollar amounts. These donations generally appear to have checked the box for "Land" or "Other Real Estate" instead of "Qualified Conservation Contribution" on Form 8283.

Table 1 provides summary information on deductions for conservation easements taken by individual taxpayers over the tax filing period from 2010 to 2012. On average over this time period, taxpayers claimed charitable deductions of approximately $1 billion of easements each year on about 2,500 donations. About 5 percent of transactions were in excess of $1.34 million.

TABLE 1. Average Annual Statistics on Easements (2010–2012)

Total deductions	$1,052,103,000
Number of individual easement deductions	2,461
Average deduction	$427,500
Median deduction	$101,250
25th percentile deduction	$43,750
75th percentile deduction	$242,000
95th percentile deduction	$1,340,000
Reported acreage?	34%
Average acreage reported	245
Median acreage	80
Mean deduction/acre	$14,750
Median deduction/acre	$1,600
Sample N (unweighted)	863

NOTE: Dollar amounts in real 2016 dollars.
SOURCE: Office of Tax Analysis, U.S. Department of the Treasury.

About 34 percent of deductions reported acreage; some deductions do not include descriptions of the property and many easements are on properties like historical buildings, where acreage is not relevant. Among those reporting acreage on Form 8283, the average size of the property under easement was 245 acres; half were larger than 80 acres. In the aggregate, among properties where acreage was reported, the average deduction

[3] Form 8283, *Noncash Charitable Contributions*, is used to report noncash contributions and is generally required of taxpayers whose noncash contributions exceed $500.

per acre was about $14,750. The median deduction was only about $1,600 per acre, suggesting that while most deductions are relatively modest, the average is skewed by some transactions with large deductions per acre.

The Concentration of Value of Easement Deductions

A small amount of acreage and a small number of large donations account for most of the tax expenditure. This is illustrated in Table 2, which compares the cumulative value of deductions claimed by taxpayers (columns 2 and 4) to the corresponding cumulative share of deductions (column 1) and cumulative share of acreage (column 3).

TABLE 2. The Concentration of Contributions of Conservation Easements in a Small Number of Transactions and Acreage

Cumulative Deductions by Donation		Cumulative Deductions by Acreage*	
Fraction of Total Donations	Fraction of Total Deductions	Fraction of Total Acres	Fraction of Total Deductions
Top 2%	43%	Top 2%	26%
Top 5%	55%	Top 5%	56%
Top 10%	70%	Top 10%	69%
Top 25%	86%	Top 25%	89%
Top 50%	95%	Top 50%	96%
Top 75%	99%	Top 75%	99%

* For donations reporting acreage.
SOURCE: Office of Tax Analysis, U.S. Department of the Treasury.

The first two columns in Table 2 show that the top 2 percent of donations—roughly the largest 50 each year valued at more than $5 million each—account for about 43 percent of the total aggregate value of donations claimed by taxpayers. The top 10 percent of donations (almost 250 donations per year valued at greater than $900,000) account for about 70 percent. Hence, the tax expenditure for contributions of conservation easements is highly concentrated in a relatively small number of transactions with very large values.

Similarly, among properties that include the acreage involved, about 26 percent of deductions go to the highest deduction-per-acre (2 percent of the properties) and 69 percent to the top 10 percent. Half of the acreage—the lowest-valued—accounts for only 4 percent of total deductions. Donations that fall within the highest value-per-acre 10 percent of properties generally are valued in excess of $6,000 per acre and can reach into hundreds of thousands of dollars per acre.

Geographic Concentration

Table 3 presents the geographic concentration of deductions for easements by taxpayers based on their State of residence aggregated over the period from 2010 to 2012. Surprisingly, in these 3 years, about 36 percent of all deductions for easements were claimed by taxpayers in Georgia. California and Connecticut are the second largest beneficiaries of deductions for contributions of easements, representing about 11 percent and 7 percent respectively.

In contrast, the number of land trusts, the number of acres under easement, and total number of acres conserved by land trusts (through any means) shows a very different geographic pattern, according to the Land Trust Alliance "2010 National Land Trust Census Report" (Land Trust Alliance (2011)). For instance, Georgia is home to only 1.3 percent of the nation's land trusts, only 2.5 percent of the nation's total acreage under easement, and only 1.5 percent of total acres conserved by land trusts. California, by comparison, has roughly 10 times as many trusts and 10 times as many total acres conserved, despite the fact that Georgia taxpayers account for 3 times the total value of tax deductions. Moreover, several States that are national leaders in the number of acres under easement or acres conserved, like Maine, Montana, New Mexico, New Hampshire, Wyoming, Arizona, or Washington, receive only a de minimis share of the tax expenditure for conservation

easements. While Georgia does have State level tax credits for easements that make contributions more favorable, it is one of several States with such incentives. Several donee organizations and businesses in the southeast U.S. appear to have developed investment strategies targeted to real estate development using easements that result in very large deductions.

TABLE 3. Geographic Concentration of Easement Deductions by Residence of Taxpayers (2010–2012)

Easement Deductions (OTA)			Land Trust Census (LTA)		
Rank	State	Percent of National Total	Share of Land Trusts	Share of Acres Under Easement	Total Acres Conserved
1	GA	36%	1.3%	2.5%	1.5%
2	CA	11%	11.6%	7.4%	14.3%
3	CT	7%	8.1%	0.4%	0.6%
4	NY	6%	5.7%	3.2%	6.1%
5	PA	6%	6.1%	2.1%	3.1%
6	VA	4%	2.1%	7.3%	7.0%
7	NC	3%	2.1%	1.7%	2.1%
8	MD	3%	3.3%	1.9%	1.2%
9	TX	2%	2.1%	3.4%	2.3%
10	SC	2%	1.4%	2.4%	1.5%
11	MA	2%	9.4%	0.9%	2.1%
12	FL	2%	1.9%	0.5%	1.1%
13	WI	2%	3.4%	0.8%	0.7%
14	CO	2%	2.2%	12.9%	7.6%
15	TN	1%	1.2%	0.7%	0.9%
16	DC	1%	0.2%	0.0%	0.0%
17	RI	1%	2.8%	0.1%	0.2%
18	AL	1%	0.5%	1.1%	1.0%
19	NV	1%	0.3%	0.0%	0.3%
20	VT	1%	2.1%	5.2%	3.8%
Total (Top 20 States)		**94%**	**67.6%**	**54.7%**	**57.5%**

SOURCE: Office of Tax Analysis, U.S. Department of the Treasury, and Land Trust Alliance, 2010.

Donee Organizations

The IRS SOI Division produces an annual public use microdata file of a sample of tax-exempt organizations drawn from the population of return entities that file Form 990 (IRS 2011).[4] The most recent file available at the time of writing is for 2011. While this is a sample and does not include all charitable organizations, large organizations (defined by assets and/or income) are sampled with certainty. Hence, starting with the IRS public use microdata file and selecting organizations that report holding conservation easements on Form 990 Schedule D, it is feasible to construct a sample of charitable organizations that report receiving easements.

Governmental organizations—Federal, State, or local governments—are also eligible to receive donations of easements. However, these organizations do not report the number, value, or characteristics of conservation easements they receive to the IRS (or to any other publicly available repository). Hence, those entities are excluded from the IRS sample.

One important caveat when examining these data is that many conservation organizations do not report donations of conservation easements on their annual returns and, if they do, many report a nominal value such as $1 for the value of the easement. From a sampling perspective, because the stratified sampling is based on the value of the organization's assets, organizations that assign a nominal or zero value to their easements

[4] In addition to the information in the IRS file, returns of all exempt organizations are made publicly available. Individual reports are thus available online to registered users (e.g., through www.guidestar.com).

are less likely to be sampled, especially if they have few other assets or otherwise concentrate in preservation using easements. The organizations that are included in the stratified sample appear to provide more complete accounting for their noncash contributions and also appear more likely to receive grants or gifts of cash. Hence, the organizations that appear in the sample may be less representative of the true population in the sense that they are likely to have greater public support in the form of cash donations and thus to have more substantial resources to devote to operations, employment, and maintenance of their easements.

This sampling convention turns out to result in a particularly severe bias against including land trusts that specialize in conservation easements, particularly easements with very large appraised values. As discussed later, a majority of the largest land trusts (measured by total deductible charitable gifts received) do not appear in the SOI public use microdata file because they have few or no other noneasement contributions and report few if any assets, and thus are not included in the sample.

Nevertheless, these data provide the best available public sample of the types of organizations that receive conservation easements, the amounts they receive, the number and acreage of the easements they hold, and the time and expense they devote to monitoring easements. However, the value of contributions of easements is not recorded in these data. To augment these data, we used the text of Form 990 available from www.guidestar.com to gather information on whether the value of contributions of easements received was recorded and, if so, the value of easements received by the organization on each of the latest 3 years of forms publicly available (Guidestar 2016). In particular, from the sample of organizations that report holding easements, we examined the 990s of the 50 largest organizations ranked by total gifts, the 10 organizations that report the most easements that year on Schedule M (noncash contributions), and the 10 organizations that report holding the most individual easements on Schedule D. We aggregated the total dollar value of easement donations received on the last three public tax returns (Form 990) filed by these organizations as of 2015.[5]

Table 4 presents this sample sorted first by the average annual value of deductions of easements received (column 2) and second, for organizations that do not report easements, by the total number of easements held. Columns 3 through 8 report information reported on their returns in 2011 including total gifts reported (including cash and, when reported, the value of noncash contributions), the number of employees, the total number of easements held, the total acreage of easements, and measures of the time and expense incurred in maintaining and enforcing the easements they manage.

The table illustrates substantial variation between the total value of easement donations received—the tax expenditure—and the size and conservation effort provided by the entity. For instance, among the organizations for which the value of easements is publicly available, the second largest recipient of such donations over the period from 2010–2013 was the Nature Conservancy, which maintains almost 2.9 million acres, has 3,725 employees, and spends more than a million dollars each year maintaining and enforcing their easements. In 2011, for example, according to their Form 990 Schedule M filing, they received 76 easements valued at $95 million total (about $1.25 million each).

In contrast, the recipient of the largest reported total value of donations of easements over the sample period was the Foothills Land Conservancy of Maryville, TN, which reported only four employees and spent only $19,000 in 2011 to monitor the 19,600 acres of easements it maintains. Over the 3-year period from 2011 to 2013, it received an annual average of $125 million in easements. In 2013 alone, it received 14 contributions of easements valued at $236.7 million (about $17 million each). (In its more recently available 2014 return, it reported an additional $206 million). For perspective, the $236.7 million in gifts in 2013 is roughly a quarter of the total volume of deductible donations of qualified conservation easements claimed by individual taxpayers that year ($1.1 billion). Indeed, its total charitable contributions places the organization alongside America's largest public charities. For instance, the $206 million in donations it received in 2014 ranked it at the 67th largest charity according to Forbes's list of the largest U.S. charities for 2016, just ahead of the Smithsonian Institution, the March of Dimes Foundation, the Humane Society of The United States, and the USO (Forbes 2016).

[5] Because of the lag between when returns are filed and when they are published, and because some organizations use different fiscal years, the 990 data refer to returns filed between 2010 to 2013. In particular, returns are available from 2010, 2011, and 2012 for some organizations, but for 2011, 2012, and 2013 for others.

TABLE 4. Organizations Receiving Conservation Easements in IRS Public Use Microdata Sample 2011 and Form 990

Name of Organization	Avg. Annual Value of Conservation Easements Received (Last 3 990s filed by 2015)	Total Gifts (2011 $)	# of Employees (2011)	Total Number of Easements (2011)	Total Acreage of Easements (2011)	Hours Spent on Enforcement (2011)	Expenditures on Enforcement (2011 $)
Foothills Land Conservancy	$866,000	40,353,481	4	96	19,638	300	19,133
Nature Conservancy	$789,000	530,700,000	3,725	2,367	2,888,283	22,496	1,065,954
The Trust For Public Land	$748,000	66,141,921	378	13	1,448	92	5,731
The Conservation Fund: A Nonprofit Corp.	$542,000	60,547,309	157	32	118,362	-	-
Rocky Mountain Elk Foundation Inc	$401,000	29,548,595	151	179	24	14	107,031
Peconic Land Trust Inc	$334,000	48,994,145	38	107	2,607	540	38,624
Natural Lands Trust Inc	$283,000	9,455,374	82	222	18,001	1,190	51,441
Wetlands America Trust Inc	$280,000	23,861,578	0	427	366,705	3,812	227,408
Triangle Land Conservancy Inc	$47,000	4,130,659	15	69	5,906	281	7,625
Little Traverse Conservancy Inc	$8,000	9,182,116	14	214	20,735	3,050	145,000
Puerto Rico Conservation Trust Fund	$866,000	16,242,950	132	5	74	62	19,748
Save The Redwoods League	$789,000	8,313,408	39	26	14,240	728	121,824
Upper Savannah Land Trust	$748,000	2,783,500	0	53	30,571	347	-
Sheriff's Meadow Foundation	$542,000	914,590	9	41	858	202	5,226
Iowa Natural Heritage Foundation	$401,000	2,197,943	40	109	14,874	892	6,163
Columbia Land Trust	$334,000	7,048,797	29	35	1,055	506	15,089
National Audubon Society Inc	$283,000	2,588,879	1,059	27	3,835	113	6,399
Open Space Conservancy Inc	$280,000	7,856,354	0	272	22,761	421	20,941
Mississippi Land Trust	$47,000	58,389,766	0	94	86,156	564	30,000
Society For Protection Of New Hampshire Forests	$8,000	5,592,776	90	795	130,189	6,628	261,021
Legacy Land Conservancy	$866,000	534,303	7	64	2,596	748	28,855
Freshwater Land Trust	$789,000	543,646	8	12	1,862	225	20,000
Brandywine Conservancy Inc	not reported	1,569,075	156	441	34,180	6,194	145,716
The Trustees Of Reservations	not reported	8,138,979	695	362	20,001	5,828	216,058
Aspen Valley Land Trust	not reported	284,551	3	231	34,379	2,500	89,700
Maine Coast Heritage Trust	not reported	16,098,023	70	204	16,725	938	128,393
Essex County Greenbelt Association	not reported	1,555,012	15	204	6,084	654	18,674
Western Pennsylvania Conservancy	not reported	6,910,782	255	163	32,507	3,247	131,368
New England Forestry Foundation Inc	not reported	1,059,480	10	145	1,144,653	1,740	81,869
The Scenic Hudson Land Trust Inc	not reported	2,339,029	0	125	12,263	2,726	101,194
Napa County Land Trust	not reported	4,707,629	15	125	21	10	128,000
Historic Landmarks Fdn Of Indiana Inc	not reported	1,347,683	51	124	477	316	8,073
American Farmland Trust	not reported	5,011,333	75	113	44,188	1,534	103,046

SOURCE: IRS Form 990 Public Use Microdata and www.guidestar.com.

The first ten organizations on Table 4 report having received a total of about $346 million in donations of easements, on average, over the prior 3 years.[6] Given that the total amount of conservation easements claimed by taxpayers was $766 million in 2010 and $695 million in 2011, this suggests the donations received by these organizations represent a large share of contributions of all easements in those years.[7]

While the information available on Form 990 is intended to be comprehensive and to allow the general public to understand which organizations are benefiting from public subsidies for charitable donations, the reporting conventions used by some land trusts excludes the value of conservation easements from gifts they report. In addition, governmental entities are generally excluded from the IRS sample.

Data reported by individuals claiming donations of easements on Form 8283 provides a means to address these shortcomings and to provide an independent source of information on the characteristics of donee organizations. The analysis of these data suggests that many of the largest organizations that receive easements do not record the value of easements on their public filings.[8] Taxpayers are required to include the name of the donee on Form 8283, as well as other information about the contribution, when filing their returns. Most donee organizations listed on Form 8283 as recipients of a conservation easement are charitable land trust organizations, but several are governmental organizations.

Table 5 summarizes the information pertaining to the 100 largest organizations based on the average annual dollar value of contributions received over the period from 2010 to 2012. As the table shows, taxpayers claimed an average of $61 million in charitable deductions for contributions of easements to each of the top 5 organizations (column 2). Contributions to those five organizations are thus collectively responsible for approximately 29 percent of the value of all deductions claimed over that period (column 3). Column 4 provides the cumulative share of all contributions made to organizations in each group and higher; the data in this column show, for instance, that the top 25 organizations account for almost half of all deductions.

Outside of the top 50 donee organizations, the typical organization receives only a few donations a year (outside of the top 100, the average is on the order of one donation per year). Given that there are approximately 1,700 land trusts in the United States, this suggests that over the 3-year period centered around 2011, roughly half of the tax expenditure was concentrated within about 1 percent of organizations.

TABLE 5. Characteristics of Donee Organizations 2010–2012

Rank (by Donations Received)	Average* Annual Gifts Received per Donee	Fraction of Aggregate Deductions	Cumulative Aggregate Deductions	# Reporting Gifts on Form 990	Average* per Donation	Donations per Year
1–5	61,462	29%	29%	2	1,770	35
6–10	20,799	10%	39%	1	639	33
11–15	10,115	5%	44%	1	1,445	7
16–25	4,434	4%	48%	2	174	26
26–50	1,156	3%	51%	na	118	10
51–100	974	5%	55%	na	228	4

* Dollar amounts in $1,000 of 2016 dollars. Estimates from individual samples 2010–2012..
na=not available.
SOURCE: Office of Tax Analysis, U.S. Department of the Treasury.

A surprising finding in this analysis is that only 6 of the top 21 nongovernmental organizations report a (nonzero) value for gifts of conservation easements on either Form 990 or on Schedule M (noncash contributions).[9] Hence, of the 21 public charities that receive the most gifts of donations of easements, 15 do not report

[6] The reports pertain largely to 2010, 2011, and 2012, but for several organizations the most recent filings are 2011–2013.

[7] The comparison is imperfect because the SOI report is based on deductions claimed by taxpayers, which may not perfectly correspond to deductions reported by organizations because of differences in the timing of when returns are filed and for which tax years, and because individuals may not claim all of the value of the deduction in the year filed and instead may carry them forward for as many as 15 tax years.

[8] Because these data are drawn only from individual returns (and thus exclude corporate donations) and because some taxpayers either do not report the donee or the donee name could not be accurately transcribed, the table likely understates the concentration of donations in these organizations.

[9] Of the top 25 recipients, 4 are government agencies, which do not file Form 990.

those gifts as receipts, assets, or otherwise on their public disclosures. The Form 990 instructions require only that donee organizations report the value of gifts for tax purposes using the same methods as they use for their books. Many organizations take the view that, for financial purposes, easements have no value because they cannot be sold or otherwise monetized.[10]

A consequence of this accounting approach is that organizations that receive noncash contributions, even those that receive hundreds of millions of dollars of public support in the form of deductible contributions, avoid public oversight that the disclosure of Form 990 is intended to provide. While many still file Form 990, Schedule D (indicating that they received qualified conservation easements), and Schedule M (noncash contributions), some land trusts that receive conservation easements avoid filing either supplementary schedules because they qualify (based on gross receipts and assets) to file either the 990N or 990EZ. Several of these organizations would otherwise rank among the top 100 or even top 50 largest charitable organizations if they reported gifts of easements at their appraised value rather than at zero.

This practice may also allow donee organizations to sidestep an important legal test required to qualify as a public charity. In particular, their noncash contributions are excluded from gifts reported on Schedule A, which is used to determine whether the organization meets the public support requirements necessary to be a public charity. Given the size and concentration of certain noncash contributions, this could affect whether certain organizations qualified to be public charities or were instead private foundations. This distinction is particularly important for these entities and for their contributors, because donations of conservation easements to private foundations do not qualify for a tax deduction for the donor.

What Is Causing the Concentration of Activity?

Qualitatively, the descriptions of the donations included on Form 8283, the characteristics of donee organizations, and other public information provide some insight into why donations are so concentrated.[11] First, many large donations appear to be associated with large real estate developments, such as a recreational community surrounding a golf course or tennis club, or a suburban residential development in which multiple homes are built on a large parcel. Because the value of the donation of the easement is generally based on its "highest and best" private use, the development of recreational amenities or high-value residences increases the value of any adjacent or undeveloped parcels of the land. For instance, building roads, installing infrastructure, and landscaping an undeveloped property increases its value in the private market, which also increases the value of the charitable contribution.

The publicly available maps from the website www.conservationeasement.us provide a selection of easements that illustrate development-related easements that could potentially result in large charitable deductions (National Conservation Easement Database (2016)). Figure 1 presents one example of an easement that is integrated with a housing development. The areas in blue were the subject of a conservation easement. Because the property is in a suburban area and because the homes are likely to be valuable, the valuation of the development rights is likely to be high.

[10] It appears that other noncash contributions, such as for art or collectibles, are sometimes treated similarly.

[11] For instance, www.conservationeasement.us provides GIS-coded maps of easements superimposed over Google Maps. Within these maps, it is possible to identify easements that are closely integrated with owner-occupied housing developments or recreational facilities, where the borders of the area under easement were drawn up to the edges of the roads and residences in the developments or superimposed over recreational facilities. Because the included land is adjacent to high-value developments and located in relatively high-cost suburbs, the private value of the land is likely to be high.

FIGURE 1. Example of a Conservation Easement Within a Housing Development

Tabulations from information regarding golf course-related easements currently under audit by the IRS suggest that such developments are heavily represented among large donations. In these transactions, the owner of a golf club pledges not to build houses or otherwise develop their golf course, and to keep it in its current use (i.e., as a privately owned and operated course) for purposes of environmental conservation. The owner then takes a charitable deduction for the diminution of value of the property. The average deduction claimed for the roughly 2 dozen golf course easements currently under audit is about $19 million and individual transactions can exceed $50 million. A single such transaction can therefore be 5 to 7 percent of all donations in a year. Indeed, the Obama Administration's Fiscal Year 2017 Budget included proposals to eliminate easements associated with golf courses and "air rights." The revenue estimates associated with those proposals suggest that roughly 10 percent of all easements are associated with golf courses and 5 percent with "air rights."

Donations of Golf Courses?

Golf courses, with manicured greens, crisscrossed by paths, maintained with fertilizers and pesticides, and surrounded by condominiums and club houses might not seem like the idyllic environmental ecosystems envisioned by the conservation-minded legislators who first introduced tax benefits for conservation easements—but these properties claim a disproportionate amount of the benefits. How so?

In one of the first major cases, the elite golf course at Kiva Dunes claimed a $30-million deduction for preserving the golf course as open space—and had its claim upheld in tax court, when challenged by the IRS (Kiva Dunes Conservation, LLC v. Commissioner (2009)). The developers of Kiva Dunes had purchased the undeveloped land on a barrier island off Alabama for $1.05m in 1992 and transformed it into a gated, residential subdivision, a 141-acre golf course, and a resource community with swimming pools, tennis courts, and beach access. Subsequently, the property owner placed an easement on the golf course and donated it to the North Atlantic Land Trust, and claimed a deduction of $30,588,235.

How could a property purchased for $1m result in a $31m deduction for contributing only partial development rights? The developer argued that he could have developed an additional 370 residential lots on the property with about $170,000 each. One reason for the high value—the "access to the amenities of the adjacent Kiva Dunes subdivision, including the use of tennis courts, swimming pools, and beach walkovers…" Perversely, the reason for the increased value was the development of the land itself. This decision is credited for broadening the types of properties that satisfy a conservation purpose and supporting these valuation methods (Ruchelman and Hicks (2009)).

Kiva Dunes is certainly not the only nor the largest such donation. In 2005, Donald Trump received a $39 million deduction for contributing an easement on one of his New Jersey golf courses (Rubin (2016a)). (He also took a deduction for an easement on his Westchester home's backyard (Rubin (2016b)).) And many more golf course owners have clearly taken deductions for donations of easements on their courses—even if the dollar amounts are not public—based on land records and maps.

The Rise of Development Use and Syndications

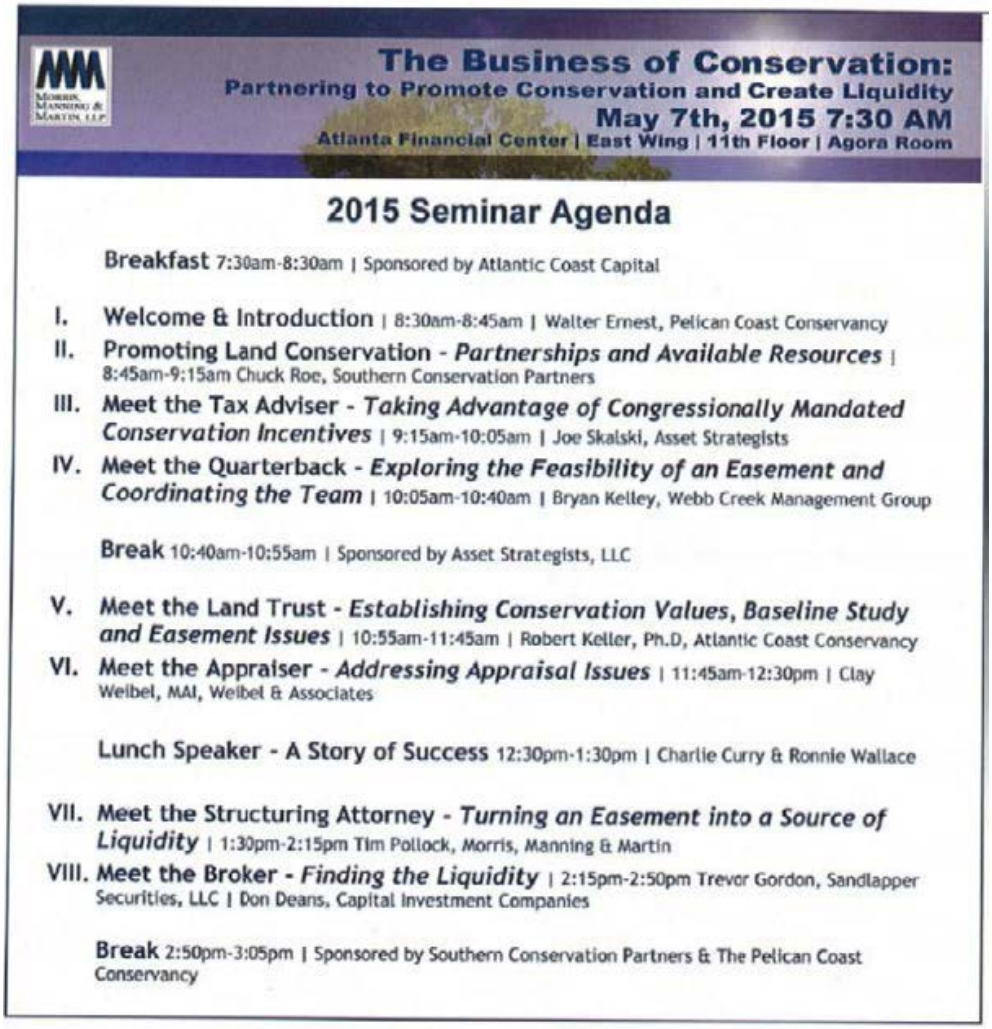
More recently, developers and other promoters have begun using syndicated transactions to expand the tax breaks they receive from easement deductions. Figures 2 and 3 provide examples of promotional materials offered to would-be investors in real estate deals that involve taking advantage of charitable deductions for conservation easements. The first document is advertised to "landowners, developers, accountants, attorneys, appraisers, land-use consultants, financial planners, and [last but not least, hopefully] wildlife resource managers" and offers seminars on topics that include "turning an easement into a source of liquidity."

The second document provides a more direct example of the promoted scheme itself in which investors are offered a substantial return in the form of a charitable deduction for investments into a real estate deal. Investors in the deal are offered the opportunity to buy one of 99 lots (just under the limit for SEC registration of an investment fund) for $36,000 and promised a return of charitable deductions of $158,000 from donation of the easement. In other words, the investor is offered a deal in which they buy the land for $36,000 but get a charitable deduction which may save them on the order of $60,000 in Federal taxes alone. In total, the 99 lots would result in $15.6 million in deductions for property apparently sold to investors for $3.6 million and presumably acquired by the current landowner for much less.

FIGURE 2. Promoted Uses of Conservation Easements in Real Estate Development

Join industry leaders to learn more about how conservation easements can be effective vehicles for land-u control and sources of liquidity. This one day seminar will offer a comprehensive overview of the d diligence associated with modern day conservation easements: tax considerations, team members and oth business dealings, documentation of conservation values, current appraisal techniques, transactic structuring pitfalls and legal defense. Space is limited. Register today to reserve your place at this event.

Who Will Benefit? Landowners, developers, accountants, attorneys, appraisers, land use consultant financial planners and wildlife resource managers will benefit from this informative seminar.



The Business of Conservation:
Partnering to Promote Conservation and Create Liquidity
May 7th, 2015 7:30 AM
Atlanta Financial Center | East Wing | 11th Floor | Agora Room

2015 Seminar Agenda

Breakfast 7:30am-8:30am | Sponsored by Atlantic Coast Capital

I. **Welcome & Introduction** | 8:30am-8:45am | Walter Ernest, Pelican Coast Conservancy
II. **Promoting Land Conservation** - ***Partnerships and Available Resources*** | 8:45am-9:15am Chuck Roe, Southern Conservation Partners
III. **Meet the Tax Adviser** - ***Taking Advantage of Congressionally Mandated Conservation Incentives*** | 9:15am-10:05am | Joe Skalski, Asset Strategists
IV. **Meet the Quarterback** - ***Exploring the Feasibility of an Easement and Coordinating the Team*** | 10:05am-10:40am | Bryan Kelley, Webb Creek Management Group

Break 10:40am-10:55am | Sponsored by Asset Strategists, LLC

V. **Meet the Land Trust** - ***Establishing Conservation Values, Baseline Study and Easement Issues*** | 10:55am-11:45am | Robert Keller, Ph.D, Atlantic Coast Conservancy
VI. **Meet the Appraiser** - ***Addressing Appraisal Issues*** | 11:45am-12:30pm | Clay Weibel, MAI, Weibel & Associates

Lunch Speaker - A Story of Success 12:30pm-1:30pm | Charlie Curry & Ronnie Wallace

VII. **Meet the Structuring Attorney** - ***Turning an Easement into a Source of Liquidity*** | 1:30pm-2:15pm Tim Pollock, Morris, Manning & Martin
VIII. **Meet the Broker** - ***Finding the Liquidity*** | 2:15pm-2:50pm Trevor Gordon, Sandlapper Securities, LLC | Don Deans, Capital Investment Companies

Break 2:50pm-3:05pm | Sponsored by Southern Conservation Partners & The Pelican Coast Conservancy

FIGURE 3. Promoted Easements Offering Substantial Returns to Investors

Subject: Conservation Easement- Possible Large Charitable Tax Deduction

Dear all, I have become aware of an opportunity to invest in a unique offering. In summary Stuckey Timberland Inc. is selling about 2,300 acres of unimproved property in Dodge County Georgia to the Parkerson Church Reserve Investments LLC (PCRI). PCRI is selling 99 units of this property at the price of $36,000/unit. Stuckey Timberland is retaining 1 unit.

If a simple majority of the PCRI unit holders vote to pursue a conservation easement then it is expected each unit holder would benefit by receiving a charitable deduction of about $158,000. In other words an outlay of $36,000 should provide a charitable deduction of $158,000. How much each investor may save on taxes from this deduction will vary. However, roughly speaking an investor with about $500,000 of income might expect a reduction of federal taxes of around $60,000. It must be noted that there is no way to generalize the value of the charitable deduction nor make any guarantees. Each investor must work with their individual CPA to estimate how much would be saved in taxes from a $158,000 charitable deduction.

Unfortunately, both time and investment units are limited. The sponsor expects this offer to be fully subscribed before 12/10/16, there are only 99 units. In order to realize the tax deduction the conservation easement must be approved and implemented before year end. This means time is short to make a decision. Please let me know by Monday 11/28 if you have an interest in this investment. I am not looking for commitment. Just interest. If there is interest we should work as a team to determine if this investment should be pursued.

That said there are risks to investing in this opportunity. There is a long list of risks delineated in the Private Placement Memorandum attached. However, some of them are:

1. The unit holders may not approve pursuing a conservation easement.

2. The company has no operating history that would permit projecting earnings. In fact there should be no net earnings if PCRI pursues a conservation easement.

3. The IRS may challenge the appraised Fair Market Value of the property and/or the intention of the RE Holding Co. to be treated as a partnership.

4. Set asides and income to pay future expenses of the property may not be enough to sustain the property.

5. There is no public market for the units and the units are subject to transfer restrictions.

6. This investment should be considered illiquid.

7. There are unlikely to be any cash distributions from the manager. All the benefits from this opportunity are likely to be derived from the charitable contribution created by the conservation easement.

Significantly, I have never been involved in a conservation easement investment opportunity, and I have no connection or history with any of the people who have created this investment. That said my preliminary due diligence suggests that the offer does seem to meet IRS guidelines, the CPA for the seller appears to have significant experience in offers like these and the same family closed a similar offer in 2015. This is the second in what is planned to be a multi-year offer.

Assuming the investment performs as expected an investor should be able to reduce their actual federal tax liability by an amount that substantially exceeds the investment cost of $36K. It appears that actual tax liability could be reduced by $50K - $70K based on the specific tax profile of each investor. However, there are risks. The investment is completely illiquid and it is possible that an investor could lose their entire investment.

Please look over the attached PPM and let me know if an investment in what is likely to be a conservation easement is of interest. If it works as expected about 2,300 acres of land will be preserved and investors should make a profit by saving taxes in excess of the amount invested.

Conclusion: What Can Be Done?

The susceptibility of the deduction for conservation easements to abuse and the acrimonious disputes arising between taxpayers and the IRS over its enforcement make the deduction ripe for reform. When used in the spirit it was intended, the deduction benefits the environment and the taxpayer by encouraging preservation of land with substantial conservation benefits. The following recommendations are intended to preserve the deduction for those purposes, while reducing the instances of abuse. The approaches outlined below try to reduce instances of abuse of the deduction by clarifying and narrowing the purposes for which a deduction can be taken; requiring organizations to meet minimum qualification standards, increasing their reporting and disclosure, and imposing minimal accountability for the transactions they facilitate; and bringing some sunshine into the area with increased public disclosure and reporting to the IRS.

1. Make Abusive Easement Transactions "Listed Transactions"

In January of 2017, the Treasury Department and the IRS identified certain easement transactions as "listed transactions" (IRS (2017)). The notice, whose purpose was to provide additional reporting of certain transactions, is important to sustain.

The notice requires participants in promoted, syndicated transactions to "raise their hand" so that the IRS and policymakers can understand the scope of these transactions and to understand whether they require greater scrutiny. The notice does not change the law or regulations related to conservation easements, and has no effect on the vast majority of traditional easement donations.

The IRS listing notice requires individual taxpayers who engaged in a specific type of transaction to provide additional information regarding their charitable contribution. There is a standard filing for taxpayers to submit this information for recent transactions and to submit in future years with their returns.

Importantly, the notice does not change the law regarding conservation easements or narrow the scope of transactions that were legally available to taxpayers before the listing notice. It is simply a tool to provide the IRS with more information on the frequency, size, and potential for abuse. Moreover, the charitable organizations themselves are specifically held harmless in the notice, so they cannot face penalty simply for being party to these transactions. And the notice does not affect the traditional donations of conservation easements.

Finally, the listing notice is extremely conservative and focuses exclusively on an extreme syndication transaction. The listing notice requires reporting only transactions with the following three characteristics: (1) the transaction was promoted; (2) it involved investments through partnership structures; and (3) the promotion offered the possibility of a charitable donation of 2.5 times or more than the investor's initial investment (i.e., "if you invest $1 million you will get a charitable deduction of $2.5 million or more.")

Since the listing notice was introduced, promoters of these abusive transactions have reacted vigorously to have the notice overturned. This is no surprise, as they are marketing and profiting from the sales of potentially billions of dollars of charitable contributions. But most charitable organizations and most conservationists find such profiteering under the veneer of environmentalism repugnant. That's why their advocates, including the Land Trust Alliance, publically support the notice.

2. Improve Reporting and Transparency

Form 990—Organization Reporting

Most charitable organizations must file annual information returns or notices with the IRS. Organizations that normally have $200,000 or more in gross receipts or assets greater than or equal to $500,000 at year-end must file Form 990. Organizations not meeting these thresholds, but normally having more than $50,000 in gross receipts, may file Form 990EZ instead. Organizations that normally have $50,000 or less in gross receipts may opt to file Form 990N ("e-postcard"), which is an electronic notice to the IRS confirming that the organization exists, providing minimal information (such as the organization's name and address). Forms 990 and 990EZ

have different "core" forms—with the Form 990EZ being shorter and easier to complete—but a variety of schedules are required to be attached to the core form if the organization is involved in specific activities or otherwise triggers filing requirements for particular schedules. For example, Schedule A, Public Charity Status and Public Support, and Schedule B, Schedule of Contributors, must be filed by Form 990EZ filers, as well as Form 990 filers, when certain threshold requirements are met.

Information regarding conservation easement contributions and holdings is required in several places on Form 990, Form 990EZ, and various schedules. Both core forms require balance sheet information, which typically does not include any value for the conservation easements held because the easements cannot typically be sold and, for financial statement purposes, are commonly not viewed by the organization as an asset, but, if anything, as a liability. Both core forms also require income statement information, which may or may not include the value of contributed easements as part of gross receipts. This is because the instructions to the core forms instruct an organization to report financial information in accordance with the way it reports information for financial accounting purposes. Some organizations report contributions of conservation easements as part of gross receipts at free-market value (FMV, using an appraised or estimated value). These organizations generally also record an expense of an equal or nearly equal amount, recognizing that after receipt of the easement, the asset will not be readily marketable and will have little, if any, value to be carried on their balance sheet. Other organizations, however, treat contributed easements as having no value for all reporting purposes, including both income and balance sheet reporting on the core forms, and for the calculation of public support on Schedule A.[12] (Note that this method of accounting is not necessarily specific to donations of easements, but may also occur with noncash gifts that effectively have no value or represent a liability to the donee, like certain restricted gifts of art or property.)

Organizations that receive or hold conservation easements that are Form 990 filers must also report information about the easements on Schedule D, Supplemental Financial Statements, and on Schedule M, Noncash Contributions. Schedule D, Part II (Conservation Easements) asks for information about conservation easements held by the organization, including the total number of easements held; the total acreage held under easement; the number of easements modified, transferred, released, extinguished, or terminated during the tax year; and the staff hours and expenditures during the tax year related to monitoring and enforcing easements. Schedule M requires reporting, by type of contributed property, of the number of contributions, the total value of the contributions and the method of valuation used and asks additional questions regarding the restrictions on contributed property and the acceptance/sale of contributed property. Form 990EZ filers are not required to file Schedules D or M.

In most places, the instructions for the Form 990 and related schedules indicate that the organization should report information, including contributions of conservation easements, consistently with how it reports information for its books, records, and financial statements. Thus, organizations reporting contributions of conservation easements as having no value on the core form, will also report no value for these contributions on Schedule M.[13] Note, however, that if an organization reports its conservation easement contributions as having no value, it is likely that it will also use the zero value for purposes of determining whether it meets the gross receipts and asset thresholds for filing Form 990. Thus, it is possible that an organization could have significant contributions of conservation easements, but normally less than $200,000 in other contributions, and under $500,000 in assets, qualifying it to file Form 990EZ. It is even possible that an organization with significant conservation easement contributions could have less than $50,000 in gross receipts annually and qualify to file the Form 990N.

[12] Since the Schedule B is not publicly available, it is not clear whether or how organizations are reporting contributions of conservation easements on Schedule B. However, it is likely that organizations reporting conservation easements as having no value on the core form are also not reporting those contributions on Schedule B. This is because the instructions to Schedule B specifically indicate that the organization should "report the value of any qualified conservation contributions and contributions of conservation easements listed in Part II consistently with how it reports revenue from such contributions in its books, records, and financial statements and in Form 990, Part VIII, Statement of Revenue." Thus, because the organization values conservation easements at zero for financial statement purposes, and the Schedule B requires reporting of contributors giving cash or property only if valued at more than $5000, these contributions would not be reported on a Schedule B.

[13] Schedule M instructions indicate that "[a]n organization that received qualified conservation contributions or conservation easements must report column (c) revenue consistent with how it reports revenue from such contributions in its books, records, and financial statements. The organization must also report revenue from such qualified conservation contributions and conservation easements consistently with how it reports such revenue in Form 990, Part VIII."

The current reporting regarding conservation easement contributions and holdings on the Forms 990 and 990EZ results in several shortcomings:

a. *Inconsistent reporting.* Different organizations report information about conservation easements in very different ways, making it impossible to compare organizations on the basis of their filings or to identify comparable organizations and/or outliers.

b. *Hides large donee organizations.* Because organizations may report the value of the contributions of easements as zero, it is difficult to identify which donee organizations are receiving donations of easements for which large charitable contributions have been taken.

c. *Understates conservation activity.* The Forms 990 and 990EZ are used to communicate to the public—including State and Federal policy makers, donors, academics, and the media—about the levels of conservation activity undertaken by the reporting organizations. When organizations report their conservation activities using a zero value, the level of activity is understated and it may be difficult to draw any meaningful conclusions from the reported information.

d. *Skews the calculation of public support, potentially allowing more deductible contributions.* Because donations of easements may be excluded from the calculation of public support, it is possible that public support calculations include only a small portion of the deductible contributions received by the organization. Thus, an organization receiving valuable easement contributions from a single donor or family, which otherwise would be a private foundation—and thus unable to receive deductible contributions of conservation easements—if easement contributions were included at FMV, may be a "qualified organization" if the easement contributions are valued at zero. This might be an appropriate outcome in some cases, but it may not be desirable in others. Moreover, the tax code includes a process for excluding extraordinary gifts from the public support test, which many donee organizations use when they receive a large gift (e.g. of an easement), but which is more transparent and follows a prescribed process.

Options for Reporting Revisions on Form 990

To address issues involving lack of transparency, Form 990/990EZ and/or the associated Schedules should be revised to require additional reporting of the FMV of contributed easements by the organizations receiving the conservation easements. In all options below, FMV would be defined as the value of the easement at the time of the contribution (in the hands of the donor). In addition, instructions would provide that organizations may use a reasonable good faith estimate of the FMV and do not necessarily need to obtain an appraisal. Further, instructions could provide that the FMV of the contributed conservation easement determined by the qualified appraiser that signs the donor's Form 8283 will be considered a reasonable estimate of the FMV of the easement unless the donee organization knows or has reason to know that the value is not correct.

Although a donor is not required to provide the donee organization with a copy of the appraisal for the contributed easement, a donee organization may request a copy of the completed appraisal from the donor (as is recommended as part of the Land Trust Alliance best practices and is required by many governmental entities that receive easements). The fact that the organization would have its own Form 990 tax reporting requirement would provide an additional basis for that request. If a donor refuses to provide a copy of the appraisal to the organization, the organization could make a good-faith estimate of the value. The organization could also refuse to sign the donor's Form 8283. Although a donor may take a deduction for a contribution if the Form 8283 has not been signed by the donee organization, the donor would need to attach a detailed explanation of the reason it was impossible to obtain the donee's signature to the form.

The following options for revising the reporting of conservation easements may be implemented alone or in combination, and may address some or all of the issues noted above.

In general, these options would require additional reporting of information about contributed conservation easements and their fair market values. In most cases, these options could be implemented by regulation or by modifying IRS guidance or instructions. In other cases, Section 6033 would be amended to require electronic reporting and public disclosure by donee organizations. These options focus on increasing disclosure regarding deductible contributions of easements that is sufficient for transparency and accountability

including: detailed descriptions of the subject property and the restrictions imposed on the property, the conservation purposes served by the easement, and any rights retained by the donor or related persons; the fair market value of both the easement and the full fee interest in the property at the time of the contribution; and a description of any easement modifications or actions taken to enforce the easement that were taken during the taxable year. As is the case under current law, personally identifying information regarding the donor would not be subject to public disclosure.

Option 1—Revise Schedule D reporting/require Form 990 filing

- Revise Schedule D to require reporting of the number and total value of conservation easement contributions during the tax year. Optionally, acreage covered by the contributed easements could also be reported. Revise the Schedule D instructions to require that the value be reported at the FMV of the contributed easements.
  - Specifically, Schedule D, Part II would be revised to add a new multi-part question requiring reporting for the tax year of (a) "Number of conservation easements received"; (b) "Total acreage covered by the conservation easements received"; and (c) "Fair market value of the conservation easements received."
  - Optionally, the cumulative total value of contributed easements could also be required by revising Schedule D, Part II, Line 2 to add another subline as line 2c for the "Total value of contributed conservation easements."
- Revisions would also be needed to ensure consistent reporting by all organizations receiving significant conservation easement contributions. Options include:
  - Require that all organizations that receive contributions of conservation easements file Form 990 (and, thus, Schedule D). This would require a change to the Form 990, Form 990EZ and Form 990N instructions.
  - Require organizations to use a reasonable estimate of the FMV of contributed easements when determining whether they meet the thresholds for filing a Form 990EZ or Form 990N. (See Option 5 below.) Because contributions of conservation easements are often valued at more than $50,000 (at FMV), most organizations receiving contributions would file at least the Form 990EZ, and only those receiving contributions of conservation easements valued in excess of $200,000 annually would be required to file the Form 990/Schedule D. This would require a change to the Form 990, Form 990EZ and Form 990N instructions.
  - Require that Form 990EZ filers also attach Schedule D if they receive contributions of conservation easements. Other parts of Schedule D could also be required of Form 990EZ filers, if desired. This would require a change to the Form 990EZ instructions.
  - Although a donor is not required to provide the donee organization with a copy of the appraisal of the contributed easement,[14] a donee organization may request a copy of the appraisal from the donor, with the tax reporting requirement providing a basis for that request. If a donor refuses to provide a copy of the appraisal to the organization, the organization would be able to make a good faith estimate of the value—or it may refuse to sign the Form 8283. Although a donor may take a deduction for a contribution if the Form 8283 has not been signed by the donee organization, a detailed explanation of the reason it was impossible to obtain the donee signature must be attached to the form.

By itself, this option would provide a minimum level of consistent reporting across organizations, allowing transparency as to which organizations are currently receiving tax-deductible contributions of conservation easements from all types of taxpayers. It would also provide an additional measure of the relative size of organizations' conservation programs. This information would be helpful in creating and evaluating policy alternatives, and could be useful for State and IRS enforcement efforts.

[14] Currently, a donor is required to provide the donee organization with a copy of Section B of the Form 8283 that includes only the name and SSN of the donor and a description of the property contributed. The donor is not required to provide a copy of the appraisal to the donee organization. This position is reflected in final regulations that are currently in clearance. The government could consider, if desired, later amending the regulations and instructions to Form 8283 to require the donor to provide more information, including the appraised fair market value and a copy of the appraisal, to the donor.

Option 2—Revise Schedule B reporting

- Revise the Schedule B instructions to require any reporting of conservation easement contributions to be at FMV.
- It appears from the review of the IRS forms that other property contributions, such as art and historical treasures, may also be given a zero value for financial statement purposes.[15] It would seem that the FMV of these contributed items should also be disclosed—particularly given concerns regarding inappropriate deductions for contributions of art works in media reports in recent years. However, this revision could also be limited to just conservation easements.

This option might be useful for IRS enforcement efforts, particularly if combined with Option 1. However, because the Schedule B is not publicly disclosed, this option alone would not provide much transparency into the contribution and holding of easements. Moreover, legislation has proposed eliminating Schedule B.

Option 3—Revise Schedule M reporting

- Revise the Schedule M instructions to require reporting of conservation easement contributions at the FMV of the contributed property. Also provide that if the organization reported no value for the contribution in the Statement of Revenue in the core form (Form 990, Part VIII), an explanation of why no value was reported should be included in Schedule M, Part II.
  - Specifically, the Schedule M instructions for Column (c)–(d), on page 3, Column 2 would need to be revised and the paragraph in Column 3 specifically indicating that qualified conservation contributions should be reported consistent with the organization's bookkeeping method would need to be eliminated.
  - Additionally, an example could be added to the instructions to illustrate how an organization would report the FMV of a conservation easement in Column (c) and, if desired, illustrate reporting in Part II of an explanation of valuation at zero on Form 990, Part VIII (Statement of Revenue).
  - Instructions for Lines 13-14, on page 2, would also be revised to include specific instructions for reporting the easement at FMV, including use of the appraised FMV reported on Form 8283 or in the appraisal of the property done by the appraiser who signed the Form 8283.
  - A conforming revision of the Schedule M, Part I, Column (c) title would also be needed.
- As with Option 2 (Schedule B reporting), this revision arguably should be made for all noncash contributions reported on Schedule D (including artwork).
- Like Option 1 (Schedule D reporting), revisions would also be needed to ensure consistent reporting by all organizations receiving significant conservation easement contributions.
- Like Option 1 (Schedule D reporting), this option alone would provide a minimum level of consistent reporting across organizations, allowing transparency regarding which organizations are currently receiving tax-deductible contributions of conservation easements and permitting easier aggregation of information on contributions by all types of donors (individual, corporate, etc.). However, these would be easier to make because changes are primarily to the instructions, and do not require structural changes in the form. This information would be helpful in creating and evaluating policy alternatives and also could be useful for State and IRS enforcement efforts.
- If this revision were implemented alone, it could result in a potential differential between the total amount of revenue reported in Form 990, Part VIII (which may include easement contributions reported at a zero value) and the total of the contributed property reported in Schedule M, Part I, Column (c). However, the organization would be required to identify and explain where this was the case in Part

[15] See Instructions to Schedule M, page 3, column 2 below Example 2.

II. If implemented in combination with Option 5 (Core Form reporting), then amounts reported on Schedule M would be consistent with amounts reported in the core form.

Option 4—Require reporting of contributions of conservation easements at FMV in Form 990 and Form 990EZ core forms

- Revise the instructions for Form 990 to require contributions of conservation easements to be reported at FMV as receipts, and also to record an equal amount recorded as an expense, reflecting a write-down of the value of the assets for balance sheet purposes.
    - Specifically, that would require changing the instructions in Part VIII (where it says "The organization must report any contributions of conservation easements and other qualified conservation contributions consistently with how it reports revenue from such contributions in its books, records, and financial statements."). In addition, the language pertaining to Line 1g should specifically call out donations of easements. One option is to add a sentence to say something like "if you acknowledged receiving noncash contributions on Form 8283, refer to the instructions for Schedule M for how to include their value on line 1g." Or something more specific, like "…noncash contributions of real property, art, qualified conservation contributions or other property…then fill out Schedule M." "If the value of these contributions, as determined according to the instructions on Schedule M exceed $25,000, then include that in 1g."

Option 5—Revise Schedule A reporting and calculation of public support

- Revise Schedule A instructions to require calculation of public support, for purposes of determining whether an organization is publicly supported and qualified to receive deductible contributions of conservation easements, using the FMV of contributed easements.

Arguably, the FMV of all contributions should be used in calculating public support, but if organizations have been calculating public support using a zero value for contributed easements, there could be some organizations, particularly smaller organizations with few easements contributed, that might be negatively impacted by such a change. Because of that, this type of a change would ordinarily require a change in regulation, with notice and comment, rather than just by changes in instructions.

Form 8283—Donor Reporting

Generally, donors may take a tax deduction for gifts of cash and property to governmental entities and "charitable organizations" qualifying for tax-exemption under section 501(c)(3), provided the requirements of section 170 are met. Generally, there is no charitable deduction allowed for contributions of partial interests in property, but an exception to this rule is provided in the case of a "qualified conservation contributions," including a contribution of a conservation easement to a "qualified organization." A qualified organization, which may receive deductible qualified conservation contributions, generally must be either a governmental entity or a 501(c)(3) organization that also qualifies as a "public charity" because it meets one of three public support tests.[16]

A donor's deduction for contributed property is generally the fair market value (FMV) of property contributed, less any gain that would not be a long-term capital gain if the property had been sold instead of donated. In addition, in some cases the deduction amount is also reduced by the amount of gain that would be a long-term capital gain, including if the contribution is made to a private foundation that does not qualify as a private non-operating foundation. Finally, the contribution amount must also be reduced by the value of any return benefit received by the donor (or related parties). For donations of property valued over certain dollar thresholds, the donor must meet certain substantiation and recordkeeping requirements.

Generally, no deduction is allowed for a contribution valued at $250 or more unless the donor obtains a contemporaneous written acknowledgement (CWA) containing specified information from the donee

[16] See Section 170(h)(3), defining a qualified organization as one which is described in Section 170(b)(1)(A)(v) (certain governmental entities eligible to receive deductible contributions), Section 170(b)(1)(A)(vi) (generally charitable organizations that meet one of two regulatory tests for receiving a "substantial part" of their support from the government or the general public), or Sections 501(c)(3) and 509(a)(2) (charitable organizations meeting a statutory public support test). Certain supporting organizations controlled by the foregoing may also be "qualified organizations."

organization.[17] In addition, for property valued over $500, donors must disclose certain information regarding the donated property on Form 8283, which is attached to the donor's tax return for the year of the contribution. For contributions valued under $5,000, only Section A of Form 8283 is required. For contributions valued over $5,000, the donor must obtain an appraisal of the contributed property and fill out Section B, as well, which requires both an acknowledgement of the contribution by the donee organization and a declaration by the appraiser. The donee acknowledgement must be signed by an authorized official, and must include the organization's name, address and EIN, and an acknowledgment both that the organization is an organization qualified to receive charitable contributions (i.e., described in Section 170(c)) and that the organization received the described property on the given date. (The organization must also specify if it intends to use the property for a use unrelated to its exempt purposes.) Donors must file separate forms for each piece of property (or group of like items) contributed.

Although there is currently some information gathered regarding contributions of partial interests (including conservation easements) valued at $5,000 or less in Section A of Form 8283, the same information is not currently requested for contributions of property valued in excess of $5,000.

Option—Improve donor reporting

- In order to take a deduction, a donor must provide a detailed description of the conservation purpose or purposes furthered by the contribution, including a description of the significant public benefits it will yield, and the donee organization must attest that the conservation purpose, public benefits, and fair market value of the easement reported to the IRS are accurate. Penalties would apply on organizations and organization managers that attest to values that they know (or should know) are substantially overstated or that receive contributions that do not serve an eligible conservation purpose.

3. Strengthen Standards for Donee Organizations and the Definition of "Conservation Purpose"

Donors have considerable latitude to determine whether an easement on their property furthers conservation purposes and over the appraised value of the easement, because the donor chooses both the organization holding the easement and the appraiser. While the majority of donors and easement holders act in good faith, there are no repercussions on those organizations that knowingly accept contributions of easements that are overvalued or do not further conservation purposes. Court cases over the last decade have highlighted donors who have taken large deductions for overvalued easements and for easements that allow donors to retain significant rights or that do not further important conservation purposes. For example, large deductions taken for contributions of easements preserving recreational amenities, including golf courses, surrounded by upscale, private home sites have raised concerns both that the deduction amounts claimed for such easements are excessive, and also that the conservation easement deduction is not promoting only bona fide conservation activities, as opposed to the private interests of donors. In addition, easement valuations often do not appropriately take into account existing limitations on the property or rights retained by donors. Reforms are needed to ensure that conservation tax benefits encourage important conservation activities and do not provide opportunities for abuse. The proposal would make changes to the deduction provision to reduce the likelihood that contributed easements are overvalued, to better ensure that contributed easements further bona fide conservation purposes, and to improve the administration and transparency of the deduction.

Option 1—Require minimum standards

- One proposal would strengthen standards for organizations to qualify to receive deductible contributions of conservation easements by requiring such organizations to meet minimum requirements, specified in regulations, which would be based on the experiences and best practices developed in several States and by voluntary accreditation programs. For example, the regulations could, among other things, specify that a "qualified organization" must not be related to the donor or to any person that is or has been related to the donor for at least ten years; must have sufficient assets and expertise to be reasonably

[17] See Section 170(f)(8).

able to enforce the terms of all easements it holds; and must have an approved policy for selecting, reviewing, and approving conservation easements that fulfill a conservation purpose. An organization that accepts contributions that it knows (or should know) are substantially overvalued or do not further an appropriate conservation purpose would jeopardize their status as a "qualified organization."

Option 2—Clarify definition of "conservation purpose"

- A second option would modify the definition of eligible "conservation purposes" for which deductible contributions may be made, requiring that all contributed easements further a clearly delineated Federal conservation policy (or an authorized State or tribal government policy) and yield significant public benefit. Rather than just strictly prohibiting the use of easements for golf courses or air rights, the intent of the clarifications regarding conservation purposes is to make sure a clearly delineated public purpose is served and to allow the public to monitor public charities.

Option 3—Require preapproval or review of easements in order to qualify for the deduction

- In the wake of similar abuses of State programs, some States appoint boards to pre-approve easements before any benefits are provided to the landowner. These boards assess the conservation purpose and environmental benefits associated with the land and the easement, and ensure that the appraisal and valuation is appropriate. For example, Colorado established its Conservation Easement Oversight Commission to pre-screen easement donations prior to issuing tax credits, and Virginia requires verification of the conservation value using criteria adopted by the Virginia Land Conservation Foundation for large easement donations. New Mexico requires that taxpayers apply for a certificate of eligibility from the Energy, Minerals and Natural Resources Department that affirms the conservation purpose and that the resources or areas contained in the donation are significant or important. This affirmative approval process is credited with reducing abuses and raising the environmental benefits of the State programs. A Federal authority or delegated authorities to such State boards could similarly reduce abuses, prevent disputes between taxpayers and the IRS, and increase the environmental and preservation returns on Federal tax expenditures.

4. Change the Deduction to an Allocable Credit for Conservation Contributions

A disadvantage of the options above is that they do not address the intrinsic incentives for individual donors to inflate their deductions or to take deductions for high-cost, low-conservation-value properties. Rather, those options simply provide more rules and oversight to limit such practices. In that regard, it involves more compliance and more burden on donors, conservation organizations, and the IRS.

An alternative approach is to replace the deduction with an allocated tax credit and effectively turn over to qualified conservation organizations the responsibility to "spend" those credits to secure easements. These organizations would have the appropriate incentives to husband the credits and to spend them to secure easements on high-value conservation purposes. By getting the incentives right, the need for strict rules and additional oversight is reduced, reducing the burden on donors, organizations and the IRS.

The Obama Administration's 2017 budget proposed to pilot a non-refundable credit for conservation easement contributions as an alternative to the conservation contribution deduction (i.e., donors taking the deduction would not be eligible for this credit) (Treasury (2016)). The credits of $100 million per year would be allocated by a Federal interagency board to qualified charitable organizations and governmental entities that hold and enforce conservation easements. The proposal also calls for a report to Congress from the Secretary of the Treasury in collaboration with the Secretaries of Agriculture and the Interior on the relative merits of the conservation credit and the deduction for conservation contributions, including an assessment of the conservation benefits and costs of both tax benefits.

Summary of the Credit Approach

The proposal would provide for a credit to be taken in lieu of the current deduction for contributions of conservation easements, incorporating and making permanent the enhanced incentives for contributions of easements. The charitable organizations and governmental entities to which the credits would be allocated

would in turn allocate the credits to those who have donated conservation easements to them that they have determined to have conservation benefit. Donors could receive credits of up to a maximum of 50 percent of the fair market value of the contributed easement and could use the credits to offset up to 100 percent of their income tax liability. Any unused credit amounts could be carried forward for up to 15 years. A revenue-neutral replacement of the deduction with credits would provide for roughly $600 million in credits per year.

This approach entrusts qualified conservation organizations with the decision-making and monitoring authority to select easements with the greatest conservation value, and provides the incentives to pay an appropriate price. Donors would have enhanced incentives to contribute because of the enhanced incentives and because the credit amount would not be limited by the donor's marginal tax rate. Finally, this approach improves administration by giving the primary responsibility for selecting and prioritizing conservation projects to the private non-profit organizations, giving it in the first instance to the interagency group that allocates the credits to qualified easement holders and, indirectly, to the qualified easement holders, who have deep knowledge of the conservation priorities and values in the communities in which they operate. The IRS would retain the ability to monitor the process through possible reporting requirements or, in extreme cases, audits of the allocating organizations.

Proposal Mechanics

(1) What organizations would be eligible to receive a conservation credit allocation?

The baseline proposal would retain the current-law definition of "qualified organizations" described in section 170(h)(3) currently eligible for the charitable contribution deduction. These organizations include governmental units and certain domestic charitable organizations.

Additional minimum standards for qualification could also be required, following the examples of States like Colorado, Delaware, Maryland, and Pennsylvania, or the accreditation program requirements developed by the Land Trust Alliance.[18] Qualification could require minimum standards regarding conservation organizations; the processes by which organizations select, review, and approve conservation easements; the management of organizations' conservation easements; and organizations' finances and governance. Additionally, in recent years there has been a move toward self-regulation by the land trust community which could also form the basis for additional qualifications for qualified easement holders.[19]

(2) How would a donor of a conservation easement receive the benefit of the tax credit?

In order to claim the credit, a taxpayer would contribute a conservation easement to a qualified easement organization that has received a conservation credit allocation. If the qualified easement holder is willing to assign a portion of its credit allocation to the taxpayer/donor, it would report to both the taxpayer and the IRS the allocation of the conservation credit to the taxpayer, including information on the donor, the property, the value of the easement, and the amount of the qualified easement holder's conservation credit allocation being assigned to the donor. The donor would be required to attach a copy of the certificate to his tax return.

(3) What Federal agency or agencies would have oversight responsibilities, including determining which organizations receive an allocation?

Oversight for qualifying organizations and for allocating credits would be a joint responsibility of the relevant Federal land management agencies, such as Department of Interior (DOI) and U.S. Department of Agriculture (USDA). In recent years, the nine bureaus and agencies that are involved in conservation efforts—particularly the lead agencies, DOI and USDA—have begun to work together to coordinate their conservation

[18] According to the Delaware Ag Conservation director, most States with farmland conservation programs are in the mid-Atlantic and include CT, DE, MD, NJ, PA, and VT. In DE, the State appraises the easement value of 100–125 farms per year. The farmer has the option of signing a contract for at least 10-year deferral of development to become eligible (Phase 1). Starting the next year, the farmer can apply for the program for the foundation (mostly State and USDA money) to buy the development easement rights. In a Dutch auction, the farmers bid an acceptable discount from FMV. The foundation buys the rights of the farms with the biggest percentage discounts up to the annual budgeted amount (typically winners bid a 60–65 percent discount). He was not sure how the other States work. He said some counties have programs: Lancaster County, PA, and Montgomery and Baltimore Counties in MD. If farmers sell at a discount, they could claim a charitable deduction for the discounted amount, but need a new appraisal due to IRS rules about appraisals within 60–90 days of settlement as this program takes a year to complete.

[19] For example, the Land Trust Alliance has created and administers an accreditation program for land trust organizations. See http://www.landtrustaccreditation.org/.

activities to increase the impact of their work. These organizations could be jointly responsible for reviewing applications of qualified easement holders and allocating the credits to the organizations based on more informed and conservation-oriented criteria, such as the capacity of the organization to hold and administer a conservation easement program and the strength and experience of their conservation strategy. These organizations would not review the easements themselves but only monitor and allocate credits among qualified organizations. The IRS would administer the conservation credit only at the donor level and only to verify credits were claimed legitimately.

(4) How would qualified easement holders apply for a credit allocation?

Qualified easement holders would apply annually for a conservation credit allocation amount, much in the same way that qualified "Community Development Entities" apply for an allocation of the New Markets Tax Credit. In their applications, the qualified easement organizations would provide information regarding their conservation easement programs and their ability to monitor and enforce conservation easements. Credits would be allocated on the strength of their conservation strategies, demonstrated records of success, and their capacity to receive and monitor easements. A qualified easement holder that receives an allocation would need to assign its allocation to taxpayers within a specified time period (e.g., 3 years) or return the unassigned amounts to be reallocated.

References

Forbes (2016). "The Largest U.S. Charities for 2016." https://www.forbes.com/sites/williampbarrett/2016/12/14/the-largest-u-s-charities-for-2016/#136678084abb.

Guidestar (2016). www.guidestar.org.

IRS (2011). "2011 Charities and Other Tax-Exempt Organizations Microdata Files." Form 990 Main Data File. https://www.irs.gov/uac/soi-tax-stats-2011-charities-and-other-tax-exempt-organizations-microdata-files.

IRS (2016). "Individual Noncash Contributions, 2013." *SOI Bulletin:* Summer 2016. https://www.irs.gov/pub/irs-soi/soi-a-innc-id1611.pdf.

IRS (2017). "Listing Notice--Syndicated Conservation Easement Transactions." https://www.irs.gov/pub/irs-drop/n-17-10.pdf.

Kiva Dunes Conservation, LLC v. Commissioner (2009). T.C. Memo. 2009-145. http://www.ustaxcourt.gov/InOpHistoric/KIVADUNES.TCM.WPD.pdf.

Land Trust Alliance (2011). "The 2010 National Land Trust Census Report."

National Conservation Easement Database (2016). www.conservationeasement.us.

National Taxpayer Advocate (2015). "Most Litigated Issues." Annual Report to Congress 2015. http://taxpayeradvocate.irs.gov/Media/Default/Documents/2015ARC/ARC15_Volume1_Most-Litigated-Issues.pdf.

Rubin, Richard (2016a). "Donald Trump's Donations Put Him in Line for Conservation Tax Breaks." *The Wall Street Journal.* March 10, 2016. https://www.wsj.com/articles/donald-trumps-land-donations-put-him-in-line-for-conservation-tax-breaks-1457656717.

Rubin, Richard (2016b). "Donald Trump Got a Big Break on 2005 Taxes." *The Wall Street Journal.* March 17, 2016. http://www.wsj.com/articles/donald-trump-got-a-big-break-on-2005-taxes-1458249902.

Ruchelman, Charles M., and Matthew C. Hicks (2009). "Kiva Dunes and Golf Course Conservation Easements: Important Implications for Tax Deductibility of Conservation Easement Contributions." http://www.capdale.com/files/Publication/DE3FB303-A38F-4010-8520-029088ADA6FC/Presentation/PublicationAttachment/5606F3DD-A54C-4517-AE82-4730AF2760D3/CMR-MH-Article%20Formatting.pdf.

Treasury (2016). "General Explanations of the Administration's Fiscal Year 2017 Revenue Proposals." https://www.treasury.gov/resource-center/tax-policy/Documents/General-Explanations-FY2017.pdf.

Tax Preparers, Refund-Anticipation Products, and EITC Noncompliance[1]

Maggie R. Jones (U.S. Census Bureau)

1. Introduction

Tax refunds are an important source of financial support and consumption smoothing for many households. The social safety net of the United States is now largely comprised of refundable tax credits that go to low-income tax filers, reflecting a transition from out-of-work to in-work benefits that dates back to welfare reform (Bitler, *et al.* (2014)). Meanwhile, the complexity of the tax system has led to the widespread use of tax preparers, especially among low-income filers who are concerned about receiving needed credits while avoiding an audit (Rothstein & Black (2015)). Preparation services are expensive, especially when the costs of refund-anticipation products are added to the tax-preparation fee. Because such products are lucrative for tax preparers, both taxpayers and preparers face perverse financial incentives to incorrectly claim tax credits or understate income on a tax filing to increase the value of a refund.

In this paper, I use a unique data set to evaluate the impact of paid preparers and refund anticipation products on Earned Income Tax Credit (EITC) noncompliance. These data are derived from the tax records and survey data, linked at the individual level, used to create the yearly take-up estimates of the EITC. The U.S. Census Bureau reports these estimates to the Internal Revenue Service (IRS). This annual report also includes an estimate of the incorrect payment rate of EITC. I am able to observe whether an incorrectly paid filer used a tax preparer and whether he or she purchased a refund-anticipation product when filing. To disentangle the paid preparer's incentives to incorrectly file from the filer's incentive, I exploit an exogenous shock to the price of refund-anticipation loans (RALs), which caused the abrupt withdrawal of tax preparers from offering this product. This shock led preparers to increase sales of a comparable product that had a lower base price, but to which preparers could add costs associated with higher-refund filings (Wu & Best (2015)). I find that this incentive led to higher rates of incorrect payment in the tax-filing season succeeding the shock—both the probability of incorrect payment and in the value of overpayment. Because of the unique nature of the data, to my knowledge this paper represents the first estimates of the direct effect of paid-preparer incentives on income tax noncompliance.

There is strong anecdotal evidence that preparers respond to the perverse incentives created by higher-priced filings, to which extra fees and products can be added.[2] Moreover, tax preparers face no regulation except for some slight oversight at the State level (Levy (2015)). However, it is extremely difficult to separate the response to incentives on the part of the preparer versus the filer. Low-income taxpayers face their own strong incentives to claim the EITC, the Child Tax Credit, and the head of household filing status (Jones & O'Hara (2016)). Taxpayers with the intent to evade may seek out tax-preparation services in the hope that using a professional will help them avoid scrutiny (Book (2009)). At the same time, filings that include many credit claims and a large refund are especially lucrative for tax preparers, who can charge a higher price for filings with multiple forms and added products.

Disentangling the response to these perverse incentives is an important task. The EITC has an overpayment price tag in the tens of billions of dollars each year—a fact that has brought intense criticism to what is otherwise widely considered an extremely effective policy (see Nichols and Rothstein (2015) for a review of

[1] Disclaimer: This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau. Research was performed under agreement TIRSE-14-M-00002 between the U.S. Census Bureau and the Internal Revenue Service.

[2] The National Consumer Law Center has multiple "mystery shopper" reports regarding tax preparation and incorrect filing (www.nclc.org).

assessments of the program).[3] Evaluating the contribution of the tax-preparation industry to the incorrect payment rate will inform policymakers of the best ways to approach reform while protecting eligible recipients.

In the next sections of the paper, I discuss the background of tax-preparation services, online filing, and the growth of refund-anticipation products. I also discuss the literature relevant to tax preparation and EITC noncompliance on the preparer side and the filer side. I then give a description of the data in section 3, focusing on the estimation of EITC eligibility and ineligibility. I present the econometric model in section 4, describing the exogenous shock and how it is used in the model. Section 5 provides the results of the econometric model, and section 6 provides some further specification analysis in support of the results. Section 7 concludes.

2. Background and Previous Literature

2.1 History of Refund-Anticipation Products

RALs came into existence during the late 1980s in tandem with the increase in electronic filing (e-filing) (Kopczuk & Pop-Eleches (2007)). They are short-term loans of a refund that tax preparers offer and consumer finance companies underwrite. In exchange for receiving refund money more quickly (often on the same day as filing), a taxpayer pays interest and fees on top of any preparation costs already incurred. Tax preparers make substantial profits on taxpayers through the use of these instruments because the loans are secured by the refund, lowering the risk of making the loan (Wu, *et al.* (2011)). With APRs generally greater than three digits, these products can be extremely lucrative. Meanwhile, a refund-anticipation check (RAC) is a temporary bank account for the taxpayer into which his or her refund is deposited. The tax preparer then writes a check or—more often—provides a pre-paid bankcard for the taxpayer's use. The taxpayer can put off paying the tax-preparation fee and have the fee taken from the refund before the balance is deposited, enabling financially constrained taxpayers to have their taxes prepared without having to pay up front. The product is essentially a lending of the tax-preparation fees, which can be substantial (Wu, *et al.* (2011)). If a taxpayer pays $30 to defer payment of a $200 tax preparation fee for 3 weeks, the APR would be equivalent to 260 percent (Wu & Feltner (2014)).

RALs were the original product offered and, until recently, had a higher price than RACs (Wu & Best (2015)). The history of RALs is tied up in the parallel history of e-filing. The IRS introduced e-filing in 1986 (Kopczuk & Pop-Eleches (2007)). E-filing had benefits both for the IRS in terms of lower administrative costs and for the taxpayer in the increased swiftness of refund receipt and the lower error rate on completed returns. The rollout of e-filing happened through a coordinated effort between the IRS and commercial tax preparers. The very first e-filing program occurred in 1985 through an IRS partnership with H&R Block. The involvement of the tax-preparation industry essentially required that taxpayers use tax-preparation services if they wanted to e-file (Davis (1998)).

Beginning in 2000, the IRS reinstated the debt indicator (which it had provided for a brief time in the 1990s). The indicator alerted tax preparers to any liens against a taxpayer's refund before they approved a RAL, which significantly lowered the risk of the loan and made RALs especially profitable. The IRS stopped providing the indicator in 2010, which led to an abrupt drop in preparers' willingness to offer the loans. The impact of the debt indicator removal occurred through the lenders who financed RALs, when the Federal Deposit Insurance Corporation notified the lenders that making the loans without the indicator would be "unsafe and unsound" (Hayashi (2016)). A simultaneous blow to the market happened in the same year, when HSBC, H&R Block's lender, was forced out of the market by the Office of the Comptroller of the Currency. For lenders who still made RALs, the size of the maximum loan they were willing to provide decreased substantially—from $10,000 to between $750 and $1,500, depending on the lender (Hayashi (2016)).

As tax preparers withdrew from the RAL market, the provision and purchase of RACs increased. Historically, RACs were less expensive than RALs, but as RACs have replaced RALs in the market, their price has gone up, bringing their price closer to that of a RAL (Wu & Best (2015)). Importantly, much of the increase in price is through add-on fees, including document processing, e-filing, and transmissions fees, which are correlated with filings for credits and thus the size of the refund. In 2011, a group of "mystery shoppers" tried

[3] http://thehill.com/policy/finance/274046-house-republicans-ask-irs-about-tool-to-prevent-improper-tax-credit-payments.

to get a sense of the accuracy and cost of tax-preparation services. The tax preparers who were studied charged the highest fees, between $330 to $540, for a tax-preparation/RAC combination for returns with qualifying-child EITC claims (Wu, *et al.* (2011)). This represents a full 20 percent of the EITC for an average recipient household (Rothstein & Black (2015)).

In 2014, the last year of data in my study, 21.6 million taxpayers obtained a RAC. The National Consumer Law Center estimates a baseline price of a RAC at $30, giving a lower bound of $648 million extracted from tax refunds just for the price of the product alone. The addition of estimated add-on fees brings this amount up to $848 million (Wu & Best (2015)).

2.2 Previous Literature

Taxpayer advocates point out that filers who use refund-anticipation products are borrowing their own money—money which will come to them in a matter of a few days if they simply e-file (Rothstein & Black (2015)). Why, then, are they so widely used? Researchers have found that consumers of these products want to get their refund money sooner and pay off more pressing debt. In a small study of taxpayers in Detroit, Barr and Dokko (2008) found that 73 percent of unbanked users of tax preparers obtained a RAL so they could pay off bills faster. In the same study, half reported that they needed to take out the loan to pay tax-preparation fees. The population of product users overlaps with users of other types of short-term, high-interest loans, including payday and title loans. Specifically, they tend to be young adults from low-income households (Feltner (2007); Theodos, *et al.* (2011)) who have children and are more likely to file as heads of household (Elliehausen (2005); Masken, *et al.* (2009)). Most users of the product have few other opportunities to obtain credit; a quarter of customers do not have a bank account or bank credit card (Elliehausen (2005)). In short, users of tax-preparation products are highly liquidity constrained and lack access to traditional banking and credit.

Providers of these products—similar to other providers of high-interest, short-term loans—use the foregoing evidence to argue that these products make consumers better off. RALs and RACs may provide the only way for low-income filers to acquire credit or to get their refund without having to pay a costly up-front fee when there are other bills to pay (Rothstein & Black (2015)). However, the economic literature on the potential benefits of short-term borrowing is mixed. While Zinman (2010) found generally positive effects of opening up credit markets to liquidity-constrained borrowers, Melzer (2011) found that access to payday lending increases a household's difficulty in paying important bills, such as mortgage, rent, and utilities. Jones (2016) found that similar access to RALs was associated with increased hardship.

In an examination of taxpayers in Illinois, Dewees and Parrish (2009) found that 60 percent of RAL users were EITC recipients, and 72 percent of RAL users in neighborhoods with a high African-American population were EITC receivers. The EITC plays an important role in the size of refunds; unlike withholdings, the EITC is refundable and offsets any tax liability (Jones & O'Hara (2016)). From a public policy perspective, much of the money that is captured from taxpayers through tax-preparation products and services are at their source transfers to low-income taxpayers from other, higher-income taxpayers (Eissa & Hoynes (2008)). When tax preparers charge usurious fees on refund-anticipation products, they capture a large portion of the safety net that is meant to support low-income wage earners (Wu & Best (2015)).

While consumer advocates object to the usurious nature of these products, a further consideration is the necessity of tax-preparation services. Filing income taxes is a federal legal requirement, and while low-income taxpayers have options to file at no cost, many may be unwilling to risk an audit or forgo needed credits through a misunderstanding of the tax system. The existence of for-profit tax preparation itself may cloud the importance of how complex the tax system is and dampen the saliency of reform (Finklestein (2007)).

An argument can be made that payments to tax preparers reduce the net benefits of the EITC for many taxpayers—especially if they would not seek out preparers in the absence of the credit. For recipients, those compliance costs would be lower if the IRS helped them prepare their returns at no cost (through volunteer tax preparation sites, for example) or determined eligibility without additional information from claimants. However, IRS administrative costs probably would be higher (unless funding on other agency activities were cut).[4] In considering the view of professional tax preparation as a substitute for IRS administrative costs,

[4] Email with Janet Holtzblatt, Congressional Budget Office, August 24, 2017.

however, thought should be given to the regressivity of preparer fees and the lack of Federal oversight of tax preparers.

These concerns are compounded by studies that indicate the opportunity of capturing public moneys through the tax-and-transfer system incentivizes fraud (Wu & Feltner (2014)). Masken, *et al.* (2009) found that taxpayers who used bank products were more often noncompliant than those who did not. When fraud is uncovered, taxpayers often bear the consequences of noncompliance; it is often difficult to prove fault on the part of the preparer, especially in cases in which the preparer has a seasonal establishment or the preparer did not sign the return (as required by law) (Levy (2015)). When the fault of noncompliance falls on the taxpayer, penalties usually include at least the reimbursement of a refund and, possibly, the denial of eligibility for credits in later tax years (Levy (2015)).

Taxpayers face strong incentives to be noncompliant, and their noncompliance may be associated with tax-preparer use. Many filers believe that using a professional preparer reduces their risk of audit (Book (2009)). There is substantial evidence that the claiming of dependents is an element of tax compliance that creates perverse incentives for filers (LaLumia & Sallee (2013); Liebman (2000); McCubbin (2000)). A taxpayer is always better off if he or she can claim a child through the combined advantages of the dependent exemption, the head-of-household filing status, the EITC, and the Child Tax Credit (conditional on meeting other eligibility requirements) (Jones & O'Hara (2016)). Thus, taxpayers who wish to get the most out of the system may stretch or violate admittedly complex tax laws, and they may provide erroneous information to preparers (Masken, *et al.* (2009)). In sum, the question of the "preparer effect" in tax fraud or incorrect credit payments is an open and complicated one.

3. Data

3.1 Data Description, Sample Description, and Summary Statistics

The data I use stem from a joint statistical contract between the U.S. Census Bureau and the IRS. The Census Bureau receives tax records from the IRS to calculate and report on the take-up rate of the EITC, with the calculation of the denominator dependent upon survey data that is representative of the U.S. population. Using survey data, I am able to determine the members of the population who appear to be eligible, regardless of whether they file a Form 1040. The earliest year of data used to compute the take-up rate is 2005, and the process of take-up calculation is reported in detail for that year by Plueger (2009). The process has changed somewhat in subsequent years, mainly in the refinement-of-income measurement.

The tax data included in the project are, for each year, Form 1040 personal income tax records, the EITC recipient file, the CP09/27 file (a record of taxpayers sent a notice from the IRS about their potential EITC eligibility), and Form W-2 records. From 2008 forward, but with the exception of tax year 2011, these records also include information on how a Form 1040 was filed (by a tax preparer, by the taxpayer online, or by the taxpayer via paper) and whether a Form 1040 filing included a refund-anticipation product (RALs and RACs are coded separately).

The survey data used in this project are yearly Current Population Survey Annual Social and Economic Supplements (CPS ASEC) from 2009 to 2015, matched at the individual level for the corresponding tax year with the IRS data (that is, 2008 to 2014). Records are linked at the Census Bureau using a process whereby individuals in each data set were given a unique, protected identification key, called a PIK. When a Social Security Number (SSN) is available in a data set (such as all of the IRS records used in this project), the identifier is assigned based on SSN. Identifier placement is close to 100 percent in the case of administrative tax records with an SSN. For records without an SSN, personally identifiable information such as name, address, and date of birth is used in probabilistic matching against a reference file to assign PIKs. Personal information is then removed from each data set before they may be used for research purposes. For the EITC estimation project, I also remove persons whose income and wage values were imputed in the CPS ASEC, as initial EITC eligibility determination is dependent on these values.

Each CPS ASEC, administered annually in March, includes questions regarding family structure and earnings that can be used to estimate tax filings for the year preceding the survey. I rely on identifiers in the

survey data to group people into tax-filing units and to establish the relationship between everyone in the household and the household reference person (also known as the "householder"). The householder is presumed to be the primary tax filer for the identified tax unit; in the case of multifamily households, family identifiers are similarly used to distinguish separate units. In a later adjustment, the tax information on eligibility is transferred from the householder to the spouse if it was the spouse who filed. For each filing unit, I then assign secondary filer status and dependents. This strategy relies on an expanded relationship variable, which reports the exact relationship of a person to the householder (spouse, parent, child, sibling, etc.). Children of the householder are further identified using a parent pointer, which designates the parent's identifier for each child in the household. Variables on tax-unit earnings, income, and dependent-support requirements (which, in the case of the EITC, is based on where a child lived for the tax year) are first taken from the survey data. Then, values from the tax data, when available, are swapped in for the survey values, and eligibility is refined based on these new values.

This strategy provides a denominator expressing everyone who is eligible for EITC in the tax year, regardless of whether they file a Form 1040. The numerator of the target take-up rate is the subset of eligible tax filers who actually file a Form 1040 and claim the credit. The incorrect payment rate is those tax filers who do not look eligible based on their modeled eligibility but who claim and are paid EITC in the tax year. The next subsection describes in detail the process of determining the incorrect payment rate and compares it to an IRS-calculated rate derived from audits. In what follows, I define the incorrect payment rate (described at length in section 3.2) as the number of CPS ASEC ineligible persons receiving EITC in a tax year divided by the number of taxpayers linked with the CPS ASEC.

The sample population is CPS ASEC respondents who file taxes. I observe EITC receipt, tax-preparation method, and refund-anticipation product purchase for this population. Also available is a rich set of filer characteristics, including age, sex, race, Hispanic origin, foreign-born status, filing status (defined as single, head of household, or married[5]), the number of dependent children claimed on Form 1040, educational attainment, adjusted gross income (AGI), investment income, self-employment status, and urban ZIP code. I also include measures of two social programs that are correlated with EITC receipt: TANF and SNAP participation.

There is always a concern, when linking data sets, that a prevalence of mismatch leads to biased estimates. In the project described in this paper, the population of interest is Form 1040 filers. There is selection into this population; however, Form 1040 filing is a precondition for each of the main variables in the analysis (tax-preparer use, product purchase, and incorrect EITC payment). Therefore, any statements about the results of the analysis should be taken to apply only to tax filers. The analysis also relies on variables from survey data; the resulting linked files for each year do not constitute the population of Form 1040 filers or the original random sample of the survey. To address this issue, I calculate the probability that a CPS respondent received a PIK[6] and use the inverse of this probability to reweight the CPS ASEC person weight and replicate weights. Because Form 1040 observations receive PIKs close to 100 percent of the time, calculating the probability that a CPS ASEC respondent receives a PIK amounts to calculating the probability that a tax-filing CPS ASEC respondent is found in the Form 1040 data. The Appendix reports on how well this strategy covers the number of Form 1040 filers in each tax year. The weighted mean and standard error for all variables used in the analysis are shown by tax year in Table 1.

The number of linked CPS ASEC persons declined over the period, both because the number of persons surveyed decreased and because there are more earnings nonresponses over time. When weighted, however, the number of linked CPS ASEC persons increased. There was a slight decrease in the rate of using a tax preparer, with about 58.8 percent using a preparer in 2008 versus 55.3 in 2014. Meanwhile, filing online with tax-preparation software increased from 32.0 percent to 40.6 percent. These numbers are in line with reports by the Director of IRS to Congress in April 2014, which reported 56 percent of returns were filed by a tax preparer and 34 percent were filed with tax preparation software.[7] The latter rate is higher in my data because I am including as "online" those who prepared their taxes using tax-preparation software, but then printed out the return and mailed it. Meanwhile, those writing out paper returns by hand and mailing them dropped from 9.3 percent to 4.2 percent. The mean of RAL use decreased from about 5.6 percent in 2008 to 0.1 percent by

[5] Married filing separately is included in the "single" category. Married persons filing separately are not eligible for EITC.

[6] The placement of PIKs for the CPS ASEC in each year is around 90 percent.

[7] https://www.finance.senate.gov/imo/media/doc/Koskinen%20Testimony.pdf.

TABLE 1. Summary Statistics for Analysis Variables by Tax Year, Weighted, 2008–2010, 2012–2014

Variable	2008		2009		2010		2012		2013		2014	
	Mean	SE	Mmean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Preparer	0.588	0.003	0.584	0.003	0.582	0.003	0.569	0.003	0.562	0.003	0.553	0.003
Online	0.320	0.002	0.336	0.003	0.358	0.003	0.380	0.002	0.391	0.003	0.406	0.003
Paper	0.093	0.002	0.079	0.001	0.060	0.001	0.051	0.001	0.047	0.001	0.042	0.001
Any product	0.138	0.002	0.139	0.002	0.138	0.002	0.142	0.002	0.127	0.002	0.130	0.002
RAL	0.056	0.001	0.048	0.001	0.013	0.001	0.001	0.000	0.000	0.000	0.001	0.000
RAC	0.081	0.001	0.091	0.002	0.125	0.002	0.141	0.002	0.127	0.002	0.130	0.002
Paid EITC	0.177	0.002	0.191	0.002	0.193	0.002	0.192	0.002	0.191	0.002	0.187	0.002
Age	44.493	0.071	45.078	0.066	45.461	0.069	45.585	0.066	45.725	0.073	45.886	0.071
Sex (Male=1)	0.651	0.002	0.651	0.002	0.648	0.002	0.637	0.002	0.635	0.002	0.634	0.002
Single filer	0.234	0.002	0.235	0.002	0.236	0.002	0.234	0.002	0.233	0.002	0.238	0.002
Head of household	0.078	0.001	0.080	0.001	0.079	0.001	0.082	0.001	0.079	0.001	0.078	0.001
Married	0.439	0.003	0.437	0.002	0.433	0.003	0.429	0.003	0.431	0.003	0.426	0.003
Number of children	0.527	0.004	0.548	0.005	0.539	0.004	0.542	0.004	0.536	0.004	0.525	0.005
White	0.819	0.002	0.818	0.002	0.814	0.002	0.801	0.002	0.802	0.002	0.798	0.002
Black	0.114	0.001	0.115	0.001	0.118	0.001	0.120	0.001	0.117	0.001	0.118	0.001
Asian	0.042	0.001	0.044	0.001	0.044	0.001	0.050	0.001	0.053	0.001	0.055	0.001
Other	0.024	0.001	0.023	0.001	0.023	0.001	0.029	0.001	0.028	0.001	0.029	0.001
Hispanic	0.135	0.002	0.134	0.001	0.134	0.002	0.145	0.002	0.148	0.002	0.151	0.002
Less than HS	0.116	0.002	0.111	0.002	0.106	0.002	0.099	0.002	0.095	0.002	0.096	0.002
HS degree	0.283	0.003	0.289	0.003	0.282	0.002	0.273	0.003	0.267	0.003	0.266	0.002
Some college	0.300	0.002	0.295	0.002	0.299	0.002	0.307	0.002	0.304	0.003	0.302	0.003
BA/BS +	0.301	0.003	0.305	0.003	0.314	0.003	0.322	0.003	0.334	0.003	0.336	0.003
No AGI	0.006	0.000	0.004	0.000	0.007	0.000	0.005	0.000	0.004	0.000	0.005	0.000
Log AGI	10.357	0.009	10.348	0.007	10.399	0.007	10.397	0.007	10.438	0.007	10.459	0.008
Log AGI squared	109.577	0.134	109.198	0.128	110.459	0.129	110.254	0.133	111.104	0.132	111.714	0.147
No TANF	0.993	0.000	0.993	0.000	0.992	0.000	0.993	0.000	0.994	0.000	0.993	0.000
Log TANF value	0.048	0.003	0.049	0.003	0.057	0.003	0.051	0.003	0.047	0.003	0.055	0.004
No SNAP	0.939	0.001	0.923	0.002	0.910	0.002	0.907	0.002	0.903	0.002	0.904	0.002
Log SNAP value	2.616	0.007	2.714	0.008	2.788	0.009	2.806	0.011	2.824	0.011	2.809	0.010
Observations	59,741		58,568		56,284		55,275		51,401		48,227	
Weighted observations	130,203,505		127,939,439		127,430,267		131,635,761		130,800,083		131,308,372	

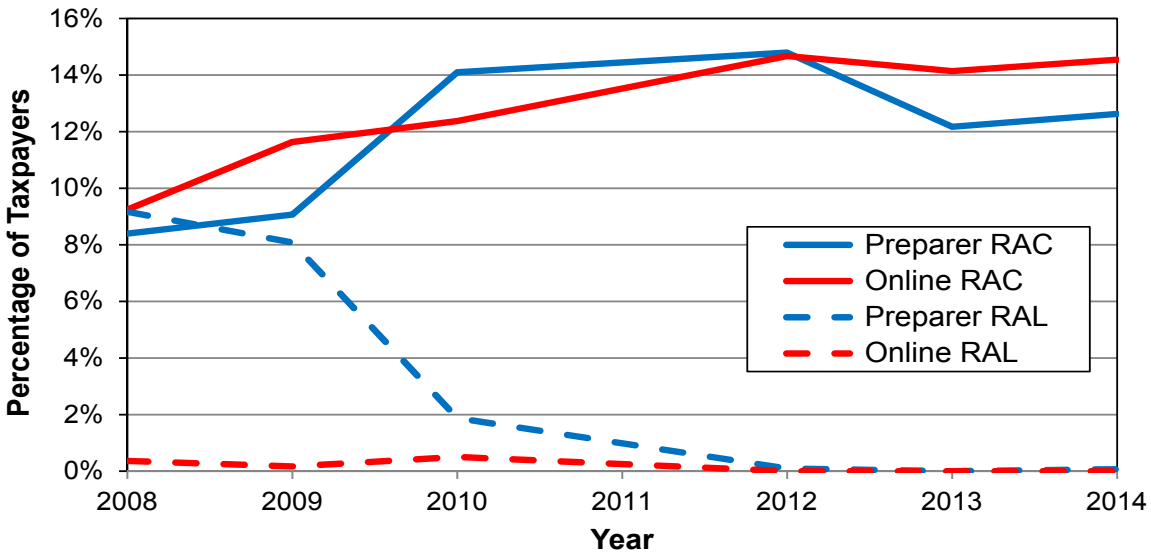
NOTE: 0.000 indicates that the values rounded to 0 but are not exactly 0.
SOURCE: Linked CPS ASEC-Form 1040 data, 2008-2010, 2012–2014.

2014, while RAC use increased from 8.1 percent to 13.0 percent. The use of any product decreased slightly: from 13.8 percent in 2008 to 13.0 percent in 2014.

Tax preparers include anyone who receives a preparer identifier from the IRS. Because there are no rules regarding licensing, a tax preparer may be a Certified Public Accountant (CPA), a volunteer preparer, an employee of a large tax-preparation company, a used car salesperson, or a travel agent.[8] Except for 2014, I am only able to identify that a tax filing was made by a preparer, and not whether the preparer was volunteer. For 2014, however, the number of taxpayers who received volunteer assistance and were linked to the CPS ASEC was fewer than the IRS disclosure threshold, indicating that the incidence of volunteer preparation is extremely low.

Figure 1 shows refund-anticipation product use over time for the matched population. At the start of the period, a roughly equal number of taxpayers purchased a RAL through a preparer or a RAC through either a preparer or online. A very small number of taxpayers purchased an online RAL. Online RACs increased over the period, while there was a precipitous drop in preparer RALs in 2010 when the IRS eliminated the debt indicator. Preparers made up for the loss in RAL use by increasing RAC sales, which experienced a slower increase than the RAL decline between 2009 and 2010. By 2012, RALs purchased through either filing method decreased to zero (or numbers fewer than the IRS disclosure threshold, which are set to 0 in the graph). The rate of RACs purchased through a preparer did not differ significantly for the rest of the period, while those purchased online continued their upward trend.

FIGURE 1. Refund-Anticipation Product Usage, 2008–2014



NOTE: RAC=refund-anticipation check; RAL=refund-anticipation loan.
SOURCE: Linked CPS ASEC-Form 1040 data, 2008-2010, 2012-2014, with year 2011 linearly interpolated.

3.2 Incorrect Payment Determination

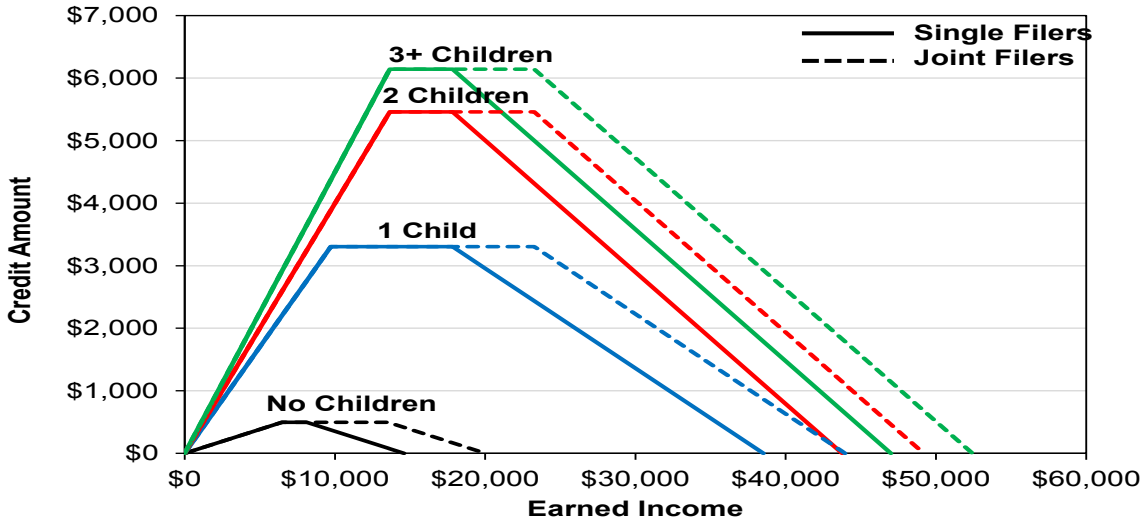
Figure 2 graphically shows the EITC eligibility and credit structure as of 2014. A filer may be precluded from eligibility in several ways besides an absence of earnings. Earning investment income more than the allowed maximum precludes eligibility. A person with either earnings or income (AGI) beyond the highest threshold for their family structure is not eligible. If income is beyond the threshold, the person is ineligible even if earnings are within range and the person claims otherwise eligible children. This category of ineligible person would be those whose combination of children, filing status (single or joint), and maximum of earnings or income are outside of the eligible range.[9] Clearly, any person with earnings or income outside the maximum

[8] Mother Jones, Secrets of the Tax Prep Business, [http://www.motherjones.com/politics/2011/04/gary-rivlin-tax-prep-refund-anticipation-loan]; e-mail correspondence with Dean Plueger, May 17, 2017.

[9] For example, a married filer with three children may have up to $52,427 in AGI compared with a single person, who must have below $46,997.

permitted in the tax year for any type of filer would be ineligible. Finally, persons without dependents are eligible for EITC only if they meet other eligibility requirements and are between the ages of 25 and 64.

FIGURE 2. The EITC Schedule for Tax Year 2014



SOURCE: Author's illustration created using program parameters outlined at irs.gov. Single filers include those who file as head of household.

A key question for the validity of the research I am presenting is whether I identify EITC recipients who are, in truth, ineligible. To be certain, I would have to require EITC recipients to undergo an audit. In the absence of an audit, I rely on comparisons between the tax data and survey data. Depending on the nature of the eligibility parameter in question, either the information reported on tax returns or the survey data is held to be the more likely measure.

The easiest determination of ineligibility occurs when there is an internal validity problem with the tax filing. For example, filers who report zero earned income on the 1040 (including Schedule C earned income or Schedule SE earned income) and have no W-2 earnings are ineligible, regardless of whether they meet other program parameters. Very few such filers claim EITC for the tax year and receive the credit.

The number of EITC eligible children claimed is a more difficult problem. The EITC conditions on child eligibility require that a qualifying child be related to the claimant biologically (son, daughter, grandchild, sibling, etc.) or be a formally adopted or foster child. A qualifying child must be either less than age 19 or less than age 24 and a full-time student, or permanently and totally disabled (at any age). Finally, the child must have lived with the claimant for more than half of the tax year. Thus the key determination in connecting potentially qualifying children to filers is that they live in the same household and should be present in the survey roster.

As outlined in Appendix B of Plueger (2009), the first step in determining which children qualify a filer for EITC is to create a qualifying-child unit identifier that collects every potentially qualifying child and every filer in the household based on survey responses. Then, the true number of children claimed for EITC by each filer on Form 1040 is compared with the claiming that was previously determined for that filer using the survey variables. For example, say filer A and filer B live in the same household (and they do not file together). If we expect, based on the survey responses, that filer A should claim a certain number of children for EITC, but he claims fewer children on Form 1040, we determine that there are "surplus" children in the household. If filer B then claims more children than he is expected to claim based on the survey responses, it seems reasonable that the "surplus" children have been appropriately claimed by filer B. When a filer claims more children than expected, I swap in Form 1040 value for the survey value only when there are surplus children in the household. Otherwise, I assume that the claiming is not appropriate, and I take the previously determined survey value for the number of children claimed.

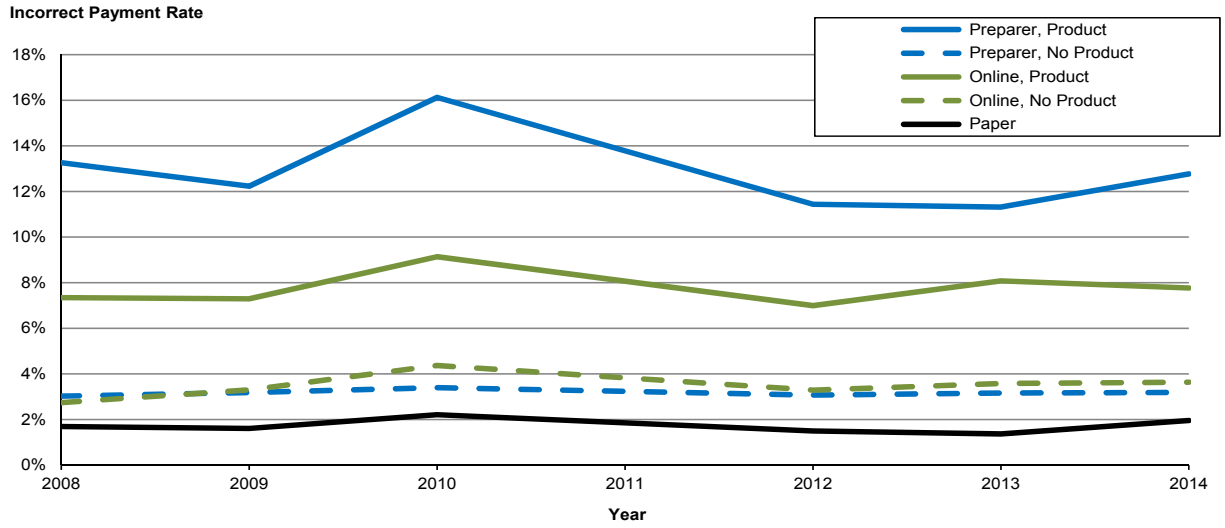
Income also undergoes adjustment based on a comparison of survey and tax data. For earnings, I swap in W-2 values for survey values when W-2 information is available for the filer. If it is not available, I swap in Form 1040 wage and salary information. W-2 information on earnings is preferred over Form 1040 information because of the well-documented problem of earnings overreporting for the EITC, mainly through self-employment earnings reports (Saez (2010)).

Finally, because filers often underreport investment income (Johns & Slemrod (2010)), I take the larger of survey-reported or Form 1040-reported investment income. Of those considered ineligible and incorrectly paid EITC due to investment income, nearly all reported investment income to the IRS that was less than their survey reports. The remainder actually claimed investment income on Form 1040, which made them ineligible for EITC.[10]

Because the population of interest for this paper is tax filers, these choices cover all of the observations in the analysis data. An instance of incorrect payment occurs when a filer receives a larger EITC credit than the estimated eligible value, with differences of $100 or less ignored. When population weights are used (reweighted to account for the probability of PIK placement), the resulting number of incorrectly paid EITC recipients in each year is close to that reported through the IRS audit process, with an incorrect payment rate for all EITC receivers of 22 to 25 percent.[11] In the Appendix, I report on categories of ineligibility and how these categories, incorrect EITC payment, and preparer and product use relate to one another.

Incorrect payment rates are intensified by the use of refund-anticipation products. Figure 3 shows how the combination of tax preparation and product is associated with incorrect payment. To avoid disclosure issues, I have combined the counts of RALs and RACs into a single product, and graphed incorrect payment for five groups defined by whether they used a tax preparer, filed online, or filed by paper, and whether they purchased a product. Those filing via paper, by definition, do not purchase a product.

FIGURE 3. Incorrect Payment Rate by Filing Type, 2008–2014



SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014, with year 2011 linearly interpolated.

The graph shows that incorrect payment rates are largely associated with refund-anticipation product use. Using a preparer and purchasing a product is associated with the highest incorrect payment rate (around 12 percent in most years, but 16 percent in 2010). Online users purchasing a product have an incorrect payment rate of about 8 percent, while both types of filers have incorrect payment rates of about 3 percent when no product is involved.

[10] The actual number who claimed investment income more than the threshold on Form 1040 was smaller than the disclosure threshold.

[11] https://www.irs.gov/pub/irs-soi/EITCComplianceStudyTY2006-2008.pdf. Here, the denominator is all tax filers rather than all EITC recipients, for an incorrect payment rate for all filers of 4.4 percent.

Incorrect payment and product use is likely codetermined by low income; the different levels of incorrect payment we see in Figure 3 may be driven by low-income taxpayers who turn to a combination of incorrect payment and product because they are liquidity constrained. Table 2 reports incorrect payment rates for each filing type, broken out by income categories. Over the period, those receiving a product from a preparer had the highest rates of incorrect payment regardless of income. Meanwhile, those receiving a product online showed lower levels of incorrect payment at each level of income. Those not receiving a product but using a preparer had incorrect payment rates between 1.6 and 6.2 percent, compared with 8.3 and 17.4 percent for the group buying a product.

The preceding look at the data provides some strong evidence that product use combined with paid preparer use has an association with incorrect EITC payment. However, the story so far cannot account for selection—perhaps taxpayers seek out paid preparation with the express purpose of claiming incorrectly and receiving quick refund money. The next section discusses an econometric model that will address the issue of selection.

TABLE 2. Rates of Incorrect Payment by Income and Filing Type, Weighted, All Years

Item	Preparer with product	Online with product	Preparer, no product	Online, no product	Paper
Less than $10K	0.083	0.049	0.025	0.024	0.012
$10K to < $20K	0.125	0.091	0.046	0.055	0.038
$20K to < $30K	0.174	0.118	0.062	0.068	0.026
$30K to < $40K	0.156	0.088	0.051	0.054	0.022
$40K to < $50K	0.154	0.086	0.048	0.048	0.017
>=$50K	0.150	0.060	0.016	0.020	0.008

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Cells show the incorrect payment rate by type of filer/product use and bins of income.

4. Econometric Model

As outlined in the preceding sections, tax preparers clearly have an incentive to sell refund-anticipation products, and they receive a higher price when the refund is larger; meanwhile, many taxpayers, based on their refund status, financial situation, and liquidity constraints, have dual incentives to buy them and to claim large refunds. Both the tax preparer's and taxpayer's incentives may become perverse vis-à-vis tax evasion or avoidance. If we see that incorrect payment occurs more often in the presence of a product and when a tax preparer is involved, how do we separate out demand effects from supply effects? Taxpayers with intent to defraud may choose to use a tax preparer in the hopes of avoiding an audit. Meanwhile, employees of tax preparation services are pushed to sell refund-anticipation products, and their bonus compensation is dependent upon these sales.[12] In the absence of some exogenous change to either supply or demand, it is impossible to disentangle these effects.

The IRS announced that it would no longer provide the debt indicator for refund-anticipation loans in August 2010, and this decision was applied to the upcoming 2011 tax-filing season. In other words, the debt-indicator removal was an unexpected move that created a price shock for the tax-preparation industry. Suddenly, this type of loan was now unsecured, which increased the marginal cost of the product in relation to its price. Banks would no longer support the loans, which meant that the tax-preparation industry withdrew quickly from the product. The industry's response to this shock can be clearly seen in Figure 1—an abrupt departure from the RAL product offering to the lower-cost RAC in 2010.

Essentially, preparers were compelled to sell RACs in 2010 to customers who previously would have purchased RALs, as shown in Figure 2. As discussed in section 2, RACs vary in price dependent on their "add-in"

[12] www.hsgac.senate.gov/download/ws-pat-eckelberry.pdf: "While I was at H&R Block, management pushed very hard to require preparers to use the client-preparer trust relationship to sell such products as Refund Anticipation Loans (RALs), Refund Anticipation Checks (RACs), and Peace of Mind (POM) guarantees.... Tax preparers received 15 percent of the system administration fee that was paid to H&R Block for completing the bank-related paperwork for a RAL. In addition, the number of RALs, RACs etc. sold was a part of our performance reviews. I have also provided a copy of my performance evaluation to the Subcommittee, which included a category for financial products such as RALs and RACs."

fees, and the highest-priced products are sold to filers with large refunds that depend on the EITC and other credits. Given little time to change the baseline price of the RAC, individual tax preparers could instead make up some of the price difference through add-on fees. To do so, preparers would have to induce a higher proportion of RAC purchasers to file for EITC. The removal of the debt indicator thus provides an opportunity to estimate a "tax preparer" effect in the incorrect payment of EITC, separate from the filer effect. In an ideal world, I would have data that would allow me to calculate the price differences between the two products from year to year. In the absence of actual prices, I use the year of the shock to detect differences in product use and incorrect payment.

I examine the impact of the removal of the debt indicator on the incorrect payment rate using a model in which I have two treatment groups (those buying a product when filing online or through a preparer); two corresponding control groups (those not buying a product and filing online or through a preparer); and a final baseline group consisting of those who file by paper (these never receive a "treatment," since "product"=0 for every member of this group). In examining the differential behavior of the group using a preparer and buying a product in 2010, I can estimate a preparer effect on the improper payment rate.

The choice of online filers as a comparison group makes the assumption that these filers are a comparable control for those who use a tax preparer with respect to the price shock in the RAL market. If the price shock induced a change in the supply of filers to online preparation, this use of the online group as the control would be suspect. However, Figure 2 provides evidence that there was no spike in online filing/product purchase in 2010. There is a steady year-by-year increase in online RAC purchases from 2009 to 2010, but the 2010 number is not statistically different from 2009. Evidence indicates that the price shock had little influence on the demand of products—filers used tax preparation and products at a similar rate but had no choice other than to accept the RAC over the RAL due to the preparers' supply decision. "Product," then, can cover either a RAL or a RAC, in that from the supplier's point of view, they differ only in their marginal cost and price structure.

The specification is as follows:

$$y_{its} = \alpha + \beta_1 prepXproductXyear_{2009} + \cdots + \beta_5 prepXproductXyear_{2014} +$$
$$\gamma_1 onlineXproductXyear_{2009} + \cdots + \gamma_5 onlineXproductXyear_{2014} + \delta prepXproduct +$$
$$\varphi onlineXproduct + \theta prep + ponline + \tau_1 year_{2009} + \cdots + \tau_5 year_{2014} + \sigma_s + X_{its}\beta + \varepsilon_{its}.$$

In this linear probability model, the dependent variable is the probability of incorrect payment of EITC. In a second specification, I also examine the value of EITC overpayment. The baseline group is paper filers without a product in 2008; dummy variables for each other year capture fixed effects over time. The variables *Prep* and *online* capture the difference in the incorrect payment rate for these groups in the absence of a product in the base year, while the interactions of these two states with *product* capture the difference in incorrect payment for these groups in 2008 when a product is used. Finally, the triple interaction with the two filing groups, product indicator, and year dummies gives the year-by-year difference in incorrect payment by group and product use. The coefficients of interest are β_2 and γ_2, the triple-interaction coefficients for 2010. The difference between them captures the impact of the debt-indicator removal, which I posit is a "preparer effect," on the incorrect payment rate. X includes the covariates reported in Table 1, and σ are State fixed effects, which are included to control for differences in State EITC, State-level lending and tax preparer regulations, and State-specific labor-market conditions. Standard errors are clustered at the State level, and the models are weighted using the recalculated CPS person weights.

5. Results

Table 3 shows the key results of the analysis. For simplicity, the table focuses on the impact of the 2010 removal of the debt indicator, reporting only the direct and interacted effects for 2010, the preparer and online dummies, and the triple interaction of these dummy variables with the product indicator.[13] Looking at model 2, which includes covariates: compared with paper filers, filers who used a preparer were about 1.1 percentage points more likely to be incorrectly paid EITC in 2008, and online filers were 1.0 percentage points more likely. The inclusion of a product in 2008 adds another 13.4 percentage points if the filer used a preparer,

[13] Appendix tables A4 and A5 report the full sets of results from this section, with the exception of the State fixed effects.

and another 5.1 percentage points if filing online. There was an increase in 2010 for those filing online and a decrease for those filing with a preparer, but the decrease for preparer filings is not statistically different from zero. There is little change between estimates from model 1 to model 2—the estimates withstand the inclusion of demographic covariates, although the coefficients on the interaction of filing type and product are slightly smaller.

TABLE 3. Results of Difference Analysis, Weighted Estimates, Dependent Variable: the Probability of Incorrect Payment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison	(4) Low income
Preparer	0.020***	0.011 ***	0.001	0.011 ***
	(0.003)	(0.002)	(0.003)	(0.003)
Online	0.018***	0.010 ***		0.012 **
	(0.003)	(0.003)		(0.004)
Preparer X Product	0.209***	0.134 ***	0.083 ***	0.129 ***
	(0.013)	(0.012)	(0.012)	(0.012)
Online X Product	0.078***	0.051 ***		0.056 ***
	(0.012)	(0.011)		(0.014)
Preparer X Product X 2010	0.060***	0.062 ***	0.033	0.061 ***
	(0.009)	(0.009)	(0.017)	(0.010)
Online X Product X 2010	0.029*	0.030 *		0.030
	(0.013)	(0.013)		(0.016)
Preparer X 2010	0.003	0.001	-0.019 ***	0.004
	(0.003)	(0.004)	(0.004)	(0.006)
Online X 2010	0.023***	0.020 ***		0.029 ***
	(0.004)	(0.005)		(0.007)
Year = 2010	0.008*	0.010 *	0.030 ***	0.012
	(0.003)	(0.004)	(0.004)	(0.006)
Product			0.050 ***	
			(0.011)	
Product X 2010			0.030 *	
			(0.013)	
State FE	yes	yes	yes	yes
Income <$60,000	no	no	no	yes
Demographic covariates	no	yes	yes	yes
Test of β2=γ2	3.23	3.61		2.30
Prob > F	0.079	0.063		0.136
Obs.	336,166	336,166	315,041	213,197

* p<0.05, ** p<0.01, *** p<0.001.

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Results from a difference-in-differences model comparing five groups: paper filers as the base category; preparer with a product; online with a product; and preparer and online with no product. Full results reported in Appendix Table A4. Standard errors clustered at the State level shown in parentheses.

The hypothesized "preparer effect" in the interplay of refund-anticipation products and incorrect payment is the difference between the triple interactions for each model. Those filing with a preparer and purchasing a product saw an additional 6.2 percentage points in the improper payment rate in 2010; meanwhile, those filing online and purchasing a product in 2010 experienced a 3.0-percentage-point increase. An F test that the coefficients are equal rejects the null at the 10-percent level. While the full model gives a better demonstration of what is going on with tax preparation and product use for all filers, we can also restrict the analysis sample to only those who file online or through a preparer. Model 3 shows the results of this exercise, which turns the previous model into a standard triple-difference. In this case, I can include the product dummy and its

interaction with each year indicator, which will capture the effect for the base group: "online." Here we can see that, compared with online filers, those filing with a preparer in 2008 had an incorrect payment rate of less than a percentage point higher, but when a product is added, the increase in rate is 8.3 percentage points. Add on the effect of filing by preparer in 2010 and purchasing a product, and the increase is an additional 3.3 percentage points, which is approximately the difference between the coefficients from the previous model 2. This estimate is again marginally statistically significant, with a *p* value of about 0.06. In this model, filers who use a preparer but do not purchase a product saw a decrease over 2008 in the incorrect payment rate, perhaps indicating that the incentive to file incorrectly decreased for clients who did not purchase a product in 2010.

Finally, I look at a low-income subset of the sample. Although I control for log income and income-squared in model 2,[14] it should be of interest to see whether effects are concentrated at the range of income where filers may be eligible for EITC along the income parameters. In contrast with expectations, the coefficients of interest on the two interaction terms are smaller and both are not statistically different from zero. An F test does not reject the null that they are the same. There does seem to be a higher likelihood of incorrect payment for this subset overall compared with the full sample, but the 2010 removal of the debt indicator did not appear to affect behavior in this subset more strongly than in the full sample.

Thus far, the modeled decision is binary: either a filer incorrectly claims EITC or does not. The next set of results examines the value of EITC overpayment and whether this varies by filing type, product purchase, preparer use, and year. Table 4 provides the mean and standard error, by year, for overpayment using all filers in the sample. Note that the largest possible overpayment value is $6,143—the maximum credit value in 2014. The dollar values are nominal (further analysis employs year fixed effects). The difference in means between online and preparer filings when a product is included is statistically significant, with preparer-plus-product filings associated with overpayments of $250 to $390 in each year (2010 is the year associated with the largest difference). Meanwhile, differences in overpayments between online and preparer filings are a few dollars and not statistically different.

TABLE 4. Weighted Means and Standard Errors of Overpayment by Year, Product Use, and Filer Type

Year	Product				No product			
	Online		Preparer		Online		Preparer	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
2008	232.68	21.41	505.07	16.50	69.42	3.65	70.03	2.84
2009	247.01	17.63	565.89	17.54	83.56	4.32	81.72	3.24
2010	305.38	19.19	694.75	22.97	116.79	4.81	91.12	3.60
2012	297.19	20.20	536.69	19.64	102.94	4.46	80.76	3.17
2013	343.96	23.77	603.43	24.31	113.57	4.92	96.09	3.88
2014	320.57	22.15	659.36	29.68	109.56	4.87	97.31	4.01

Examining this difference in the same econometric model used for estimating the probability of incorrect payment presents some challenges. Considering that incorrect payment occurs only 10 percent of the time over all taxpayers, the dependent variable in question—"overpayment"—is heavily populated with zeros. This requires a careful consideration of the choice of model. After investigation, a generalized linear model with a log link provided the most stable results and best fit. Such models are supported for the context of dependent variables with zero inflation (Nichols (2010)), especially when the zeros in question are true values and not the result of censoring. Table 5 reports the results from this analysis, using a similar difference-in-differences approach and the same set of covariates. The dependent variable is the dollar amount of overpayment, where those who were paid within $100 of their eligible amount were coded as zero. Because the value of the EITC is mechanically dependent on income, looking separately at the low-income sample is less meaningful in this case, and the model results are not reported for the low-income sample.

[14] This choice reflects the fact that over the range of the credit, value increases and then decreases in income.

The results reported in the table are coefficients; exponentiated, the coefficient on "preparer" indicates that preparer and online filings are associated with an approximate 2 percent increase in the overpayment value compared with paper filings (when all other variables are evaluated at the mean). Moreover, there is an added positive effect on overpayment for either type of filer upon product use.

For models 1 and 2, neither triple-interaction term is statistically significant. For model 3, however, the difference between preparer filings with a product and online filings with a product in 2010 is statistically meaningful. The interaction terms reported in Table 5 cannot be interpreted directly (Ai & Norton (2003)). To evaluate the marginal effect of "product" on the value of incorrect payment, I calculate each year's marginal effect (in dollar rather than log terms) for product separately for prepared filings and on-line filing, using the results of model 3 (again, all other covariates are held at their mean values). These effects are graphed in Figure 4. There is a spike in 2010 in the strength of the marginal effect for prepared filings compared with online filings, with a difference of about $125 between the two. The changes in marginal effect for prepared and online filings are statistically different from one another at the 5 percent confidence level when comparing 2008 to 2010 and 2009 to 2010 (Chi-squared statistics of 7.19 and 4.72, respectively). On the downward slope, the difference is statistically significant for the difference in changes between 2010 and 2012 (Chi-squared of 5.21). Marginal effects for both categories of filer increase over time from 2012 onward, but the differences in marginal effects are not statistically significant between 2012 and 2014.

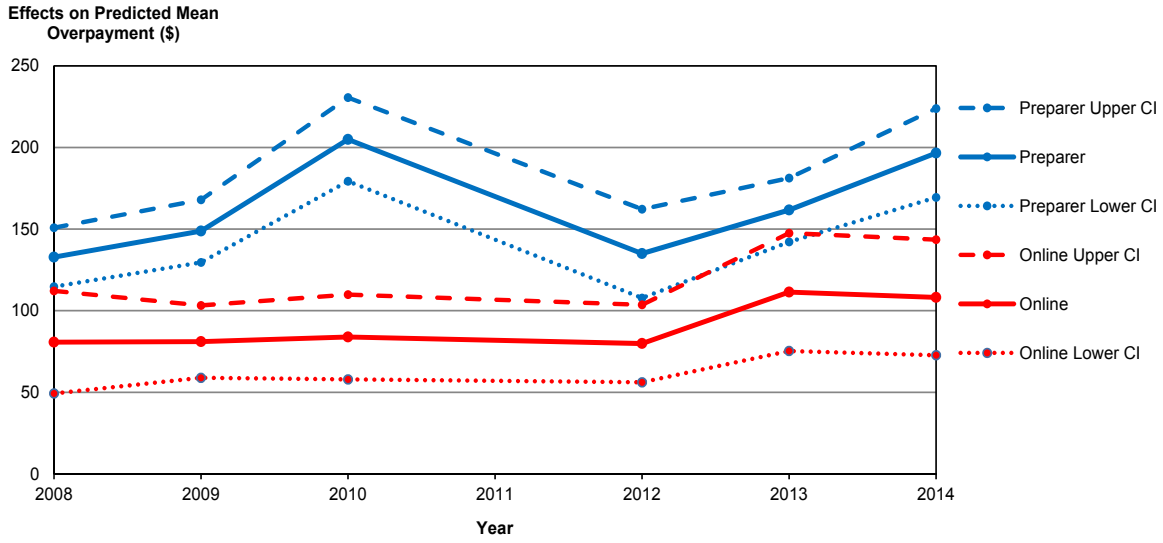
TABLE 5. Results of Difference Analysis, GLM with a Log Link, Weighted Estimates: Dependent Variable is the Probability of Incorrect Payment

Claim	(1) Baseline	(2) With covariates	(3) Online as comparison
Preparer	0.836 ***	0.663 ***	0.027
	(0.122)	(0.126)	(0.072)
Online	0.823 ***	0.641 ***	
	(0.131)	(0.131)	
Preparer X Product	1.861 ***	0.885 ***	0.171
	(0.082)	(0.055)	(0.113)
Online X Product	1.173 ***	0.714 ***	
	(0.128)	(0.122)	
Preparer X Product X 2010	0.050	0.117	0.323 *
	(0.079)	(0.072)	(0.157)
Online X Product X 2010	-0.228	-0.205	
	(0.154)	(0.138)	
Preparer X 2010	0.100	0.012	-0.247 *
	(0.166)	(0.185)	(0.097)
Online X 2010	0.351	0.258	
	(0.184)	(0.204)	
Year = 2010	0.159	0.232	0.492 ***
	(0.174)	(0.190)	(0.074)
Product			0.719 ***
			(0.122)
Product X 2010			-0.207
			(0.138)
State FE	yes	yes	yes
Demographic covariates	no	yes	yes
Obs.	336,166	336,166	315,041

* p<0.05, ** p<0.01, *** p<0.001.

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Results from a difference-in-differences model calculated using a GLM with a log link comparing five groups: paper filers as the base category; preparer with a product; online with a product; and preparer and online with no product. Full results reported in Appendix Table A5. Standard errors clustered at the State level shown in parentheses.

FIGURE 4. Average Marginal Effects of "Product," With 95 Percent Confidence Intervals



Source: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014, with year 2011 linearly interpolated.

Together, the two sets of results from Tables 3 and 5 confirm the involvement of paid preparers in the incorrect claiming of EITC, with higher overpayment opportunities inducing a greater propensity to incorrectly claim in 2010 for preparer-plus-product filings. The availability of refund-anticipation products appears to induce a response in the tax filer, whether filing online or using a preparer, but the additional propensity and value seen in the triple interaction for 2010 supports the story that preparers respond to the incentive separately.

6. Specification Checks

There are two issues that warrant investigation when it comes to the results just presented. The first is the concern that customers of preparers differed in their preference for a RAL over a RAC in relation to their desire to incorrectly claim and the value of the incorrect payment. In other words, because the preceding results depend on the treatment of RALs and RACs being such close substitutes that customers of preparers were willing to accept a RAC instead of a RAL, it is important to demonstrate that RAL and RAC customers in the pre-period did not select into the product type based on an intention to incorrectly claim. Because preparers withdrew from the RAL market, RAL versus RAC use cannot be explored as separate behavior after 2010.

To assess this issue, I examined tax filers' selection into a RAL or a RAC before 2010 based on whether the filer incorrectly claimed and the value of their EITC. The results are reported in Table 6. The sample is those filers who used a tax preparer before 2010. The dependent variable is a 1 if a filer chose a RAL, and a 0 if the filer chose a RAC. Incorrect payment is defined as before, and it is interacted with the value of the EITC, which in this case is logged. A dummy variable is included that equals 1 if a filer received no EITC (mean logged EITC replaces the continuous variable for these observations). Covariates are the same as those used in the models reported previously, and standard errors are clustered at the state level. Each of the main variables in this examination were not associated in a statistically significant way with selection into a RAL versus a RAC. All indications are that customers' selection into a RAL or a RAC did not vary in terms of incorrect claiming and the value of the incorrect claim before 2010.

TABLE 6. Results of Specification Check: Probability of Purchasing a RAL versus a RAC Before 2010

Variable	Incorrect payment	EITC amount (log)	Incorrect payment*EITC amount	No EITC
Coefficient	-0.075	0.015	0.011	0.051
SE	(0.089)	(0.008)	(0.011)	(0.060)

* p<0.05, ** p<0.01, *** p<0.001.

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2009. Results from model comparing filers who use a preparer and buying a RAL or a RAC. Number of observations: 17,363. Clustered standard errors shown in parentheses.

A second concern is whether incorrect payment of EITC was due to volume, with tax preparers handling more EITC claims that included a product in 2010 than were made online or by paper. This is examined in Table 7, which reports the same Model 2 as in the main results, but with a dependent variable coded 1 if a tax filer received EITC (regardless of eligibility).

Overall volume of EITC claimants who bought a product through a preparer did not experience a statistically significant change in 2010. While there are more EITC claimants who file via a preparer or online versus paper, and buying a product is associated with receiving EITC, there was no increase in traffic in EITC claims for the effect of interest: Both triple-interaction terms are not statistically different from zero. The results from Tables 6 and 7 indicate that the greater propensity for preparers to file incorrectly in 2010, and for credits of greater value, are not due to filers' selection into a particular kind of product or for changes in filer type specific to 2010.

TABLE 7. Results of Difference-in-Differences Analysis, Weighted Estimates, Dependent Variable = Paid EITC

Item	With covariates
Preparer	0.033 ***
	(0.005)
Online	0.032 ***
	(0.004)
Preparer X Product	0.206 ***
	(0.009)
Online X Product	0.057 ***
	(0.009)
Preparer X Product X 2010	0.005
	(0.008)
Online X Product X 2010	0.020
	(0.013)
Preparer X 2010	0.001
	(0.008)
Online X 2010	0.013
	(0.007)
Year = 2010	0.008
	(0.007)
State FE	yes
Demographic covariates	yes
Obs.	336,166

* p<0.05, ** p<0.01, *** p<0.001.

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Results from a difference-in-differences model comparing five groups: paper filers as the base category; preparer with a product; online with a product; and preparer and online with no product. Standard errors clustered at the State level shown in parentheses.

7. Conclusion

Each year, taxpayers rely on the refund money due to them from the tax and transfer system; receipt of refunds is especially important to filers who receive the EITC. Since the advent of e-filing, tax preparers have been absorbing a proportion of this refund money. Some see tax preparers as providing a valuable service, in that they can guide taxpayers through tax complexity, making sure they take advantage of the credits and exemptions they are eligible for. Yet, through high fees for preparation and—especially—for refund products, tax preparers make a substantial profit on low-income taxpayers.

These profits create perverse incentives for tax preparers to move filers into the "eligible" category when it comes to EITC. I show that rates of incorrect payment and EITC overpayment values are higher for preparer-made filings compared with online filings, and even higher when refund-anticipation products are added. But these rates may be accounted for due to choices made by the filer. Using an exogenous price shock to a particular refund-anticipation product, I estimate a "preparer effect" in the incorrect payment of EITC: in response to this shock, there was an increase in incorrect payment for filers who used a preparer and bought a product in the year of the shock. Moreover, the value of EITC overpayment was higher for preparer-plus-product filings in this year. This "preparer effect" was statistically different from the effect calculated for online filers in the same year who bought a product—marginally so in the case of any incorrect filing and at conventional levels for overpayment.

To my knowledge, this is the first paper able to directly examine the preparer effect on tax evasion (a broader issue certainly not confined to EITC recipients). The results of this paper should help inform ongoing policy discussions on how to address the incorrect payment rate of EITC while simultaneously protecting eligible receivers, as well as discussions regarding tax-preparer regulation and licensing.

References

Ai, C. & C. Norton. 2003. Interaction Terms in Logit and Probit Models. *Economics Letters,* 80(1), pp. 123–129.

Barr, M. S. & J. Dokko. 2008. Third-Party Tax Administration: The Case of Low-and Moderate-Income Households. *Journal of Empirical Legal Studies,* 5(4), p. 963–981.

Berube, A., B. Forman, M. Burns, & A. Kim, 2002. The Price of Paying Taxes: How Tax Preparation and Refund Loan Fees Erode the Benefits of the EITC, s.l.: s.n.

Bitler, M., H. Hoynes, & E. Kuka, 2014. Do In-Work Tax Credits Serve as a Safety Net?, s.l.: NBER Working Paper.

Book, L., 2009. Refund Anticipation Loans and the Tax Gap. *Stanford Law and Policy Review,* 20(85).

Costanzo, J. & A. Klekowski von Koppenfels, Counting the Uncountable: Overseas Americans, *Migration Information Source,* May 17, 2013, https://www.migrationpolicy.org/article/counting-uncountable-overseas-americans.

Davis, S. L., 1998. Unbridled Power: Inside the Secret Culture of the IRS. New York: HarperBusiness (HarperCollins).

Dewees, S. & L. Parrish, 2009. Borrowed Time: Use of Refund Anticipation Loans Among EITC Filers in Native American Communities, s.l.: First Nations Development Institute, Center for Responsible Lending, and First Nations Development Institute. Native Assets Research Center.

Eissa, N. & H. Hoynes, 2008. Redistribution and Tax Expenditures: The Earned Income Tax Credit, s.l.: s.n.

Elliehausen, G., 2005. Consumer Use of Tax Refund-Anticipation Loans, s.l.: Credit Research Center Monograph.

Feltner, T., 2007. Debt Detour: The Automobile Title Lending Industry in Illinois, Chicago, IL: Woodstock Institute and the Public Action Foundation.

Finkelstein, A., 2007. E-ztax: Tax Salience and Tax Rates, s.l.: NBER Working Paper.

Hayashi, A. T., 2016. The Effects of Refund Anticipation Loans on Tax Filing and EITC Takeup. Virginia Law and Economics Research Paper No. 2016–9.

Holt, L., 2009. Testimony by Leonard Holt, CHH Small Firms Services, at the IRS Return Preparer Review Forum. s.l.:s.n.

Holt, S., 2011. Ten Years of the EITC Movement: Making Work Pay Then and Now., s.l.: s.n.

Johns, A. & J. Slemrod, 2010. The Distribution of Income Tax Noncompliance. *National Tax Journal,* 63(3), p. 397.

Jones, M. R., 2016. A Loan by Any Other Name: How State Policies Changed Advanced Tax Refund Payments, s.l.: s.n.

Jones, M. R. & A. O'Hara, 2016. Do Doubled-up Families Minimize Household-Level Tax Burden? *National Tax Journal,* 69(3).

Kaplow, L., 1996. How Tax Complexity and Enforcement Affect the Equity and Efficiency of the Income Tax. *National Tax Journal,* 49(1), p. 135.

Kopczuk, W. & Pop-Eleches, C., 2007. Electronic Filing, Tax Preparers and Participation in the Earned Income Tax Credit. *Journal of Public Economics,* 91(7-8), pp. 1351–67.

LaLumia, S. & J. M. Sallee, 2013. The Value of Honesty: Empirical Estimates. *International Tax and Public Finance,* 20(2), p. 192–224.

Levy, A. H., 2015. Believing in Life After Loving: IRS Regulation of Tax Preparers. *Florida Tax Review,* Volume 17, pp. 437–669.

Liebman, J. B., 2000. Who Are the Ineligible EITC Recipients. *National Tax Journal,* 53(4), pp. 1165–1185.

Masken, K., M. Mazur, J. Meikle, & R. Nord, 2009. Do Products Offering Expedited Refunds Increase Income Tax Non-Compliance?. s.l., s.n.

McCubbin, J., 2000. EITC Noncompliance: The Determinants of the Misreporting of Children. *National Tax Journal,* 53(4, Part 2), pp. 1135–1164.

Melzer, B. T., 2011. The Real Costs of Credit Access: Evidence from the Payday Lending Market. *The Quarterly Journal of Economics,* 126(1), pp. 517–555.

Nichols, A., 2010. Regression for Nonnegative Skewed Dependent Variables. *BOS10 Stata Conference,* Volume 2.

Nichols, A. & J. Rothstein, 2015. The Earned Income Tax Credit. In: *Economics of Means-Tested Programs in the United States,* Vol. 1. s.l.:University of Chicago Press, pp. 137–218.

Plueger, D., 2009. Earned Income Tax Credit Participation Rate for Tax Year 2005, s.l.: s.n.

Rothstein, D. & R. Black, 2015. Improving the Tax Preparation Experience, New York: s.n.

Saez, E., 2010. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy,* 2(3), pp. 180–212.

Theodos, B. *et al.,* 2011. Who Needs Credit at Tax Time and Why: A Look at Refund Anticipation Loans and Refund Anticipation Checks, s.l.: Urban Institute.

Wu, C. C. & M. Best, 2015. Taxpayer Beware: Unregulated Tax Preparers and Tax-Time Financial Products Put Taxpayers at Risk.. In: *NCLC/CFA 2014 Refund Anticipation Loan Report.* Boston: National Consumer Law Center.

Wu, C. C., D. Del Rio, A. Iwanisziw, & P. Skillern, 2011. Tax Time 2011: Shopper Testing In New York and North Carolina Finds Continuing Problems With Tax Preparers, Boston: National Consumer Law Center.

Wu, C. C. & Feltner, T., 2014. It's a Wild World: Consumers at Risk from Tax-Time Financial Products and Unregulated Preparers.. In: *NCLC/CFA 2013 Refund Anticipation Loan Report.* Boston: National Consumer Law Center.

Zinman, J., 2010. Restricting Consumer Credit Access: Household Survey Evidence on Effects Around the Oregon Rate Cap. *Journal of Banking & Finance,* 34(3), pp. 546–556.

Appendix

1. Further Information on Ineligibility Calculation

Table A1 shows categories of eligibility and ineligibility by incorrect payment status and the numbers and percent of row in each category. First, 4.3 percent of those with zero earnings reported for tax purposes (either reported in the wage and salary field of Form 1040 or reflected in a lack of W-2 earnings) erroneously filed for and received EITC. A higher-than-average rate of incorrect payment—25.6 percent—is seen for those who reported to the CPS that they had investment income above the maximum.

TABLE A1. Eligibility/Ineligibility Categories by Incorrect Payment Status, All Years

Eligibility/Ineligibility Category	Incorrect Payment		Total
	No	Yes	
Eligible	48,503	0	48,503
Weighted observations	110,970,354	0	110,970,354
Mean	1.00	0.00	
Zero earnings	3,241	145	5,718
Weighted observations	8,496,278	381,452	8,877,731
Mean	0.957	0.043	
Investment income over threshold	771	314	1,085
Weighted observations	1,778,040	612,941	2,390,981
Mean	0.744	0.256	
Max (income, earnings) over threshold	175,327	10,560	185,887
Weighted observations	418,680,560	25,845,908	444,526,468
Mean	0.942	0.058	
Zero dependents, age out of range	92,354	1,421	93,775
Weighted observations	216,229,599	3,552,876	219,782,475
Mean	0.984	0.016	
Other eligibility loss (via model or tax status)	1,672	1,858	3,530
Weighted observations	4,265,218	4,514,564	8,779,782
Mean	0.486	0.514	
Total	321,868	14,298	336,166
Weighted observations	760,420,049	34,907,741	795,327,791
Mean	0.956	0.044	

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Cells show the number in each category of EITC ineligible tax filer and the incorrect payment rate for the type.

The cases above are straightforward—there is a clear violation of a program threshold based on a single preclusion. The largest category of ineligibles includes those whose combination of children and income lie outside the program's parameters. Filers who claimed zero dependents on Form 1040 and were outside of the eligible age ranges had an incorrect payment rate lower than the overall average, at 1.6 percent. Filers within the age range, but with a precluding combination of income and children, received incorrect payment at a rate of 5.8 percent. Finally, a small category of filers had a conflict between the survey and tax information that appeared to make them ineligible: some filers were claimed as dependents on another filer's return; others filed as single but appeared to be married according to the household roster, thus putting their total income in question. The default stance in eligibility calculation is to accept the W-2 return information and Form 1040 values as "true" unless there is strong evidence against it (such as in the case of a spouse in the CPS ASEC household roster). This last group of filers had the highest rate of incorrect payment, at 51.4 percent, but were small in number, being less than one percent of the tax-filing population.

In contrast to work by McCubbin (2000) and Liebman (2000), which indicate that incorrect payment tends to occur through confusion or fraud regarding qualifying children, it appears as though higher rates of incorrect payment dependent on the type of potential error are associated with some of the lesser-known parameters of the program, such as the investment income threshold. The overall incorrect payment rate per taxpayer is higher for paid preparers, at 4.8 percent, than for online filers (4.2 percent) and paper filers (1.7 percent).

Table A2 shows the breakdown of these errors based on whether a filer used a tax preparer, filed online, or filed a paper form, and the incorrect payment rate for each type of error. Incorrect payment rates were about average for those reporting zero earnings, with about 4.8 percent of zero earners filing with a tax preparer receiving incorrect payment, 3.3 percent of online filers, and 4.5 percent of paper filers. For those with misreports of investment income, online filers had highest rate of incorrect payment, at 36.0 percent. Interestingly, those who were ineligible based on their combination of income and qualifying children had rates of incorrect payment for each type of filer that were slightly above average, with those receiving preparation assistance at 6.6 percent and those filing online at 5.2 percent. The group with no dependents and appropriate earnings and income, but whose ages were out of range, had the lowest rates of all, at slightly less than 2.0 percent for both prepared and online filings. Finally, the "other" category had the highest incorrect payment rates, at 57.7 and 46.3 for prepared and online, respectively.

TABLE A2. Rate of Incorrect Payment by Reason for Ineligibility and Filing Method, All Years

Reason for ineligibility	Filing method		
	Prepared	Online	Paper
Zero earnings			
Incorrect payment rate	0.048	0.033	0.045
Total filers by method	1,948	1,004	434
Weighted filers by method	5,075,157	2,671,069	1,131,504
Investment income over threshold			
Incorrect payment rate	0.231	0.360	0.129
Total filers by method	749	269	67
Weighted filers by method	1,627,357	601,910	161,714
Max (income, earnings) over threshold			
Incorrect payment rate	0.066	0.052	0.017
Total filers by method	102,545	73,436	9,906
Weighted filers by method	244,352,592	176,079,508	24,094,367
Zero dependents, age out of range			
Incorrect payment rate	0.017	0.018	0.002
Total filers by method	55,424	29,455	8,895
Weighted filers by method	130,265,031	70,083,887	19,433,557
Other eligibility loss (via model or tax status)			
Incorrect payment rate	0.577	0.463	0.221
Total filers by method	2,055	1,246	229
Weighted filers by method	5,107,379	3,136,773	535,631
Total incorrect payment rate	0.048	0.042	0.017

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Cells show the number in each category of EITC ineligible filer by filing method and the incorrect payment rate for the type.

2. Reweighting and Comparison to Original Form 1040 Records

Because Form 1040 observations receive PIKs close to 100 percent of the time, calculating the probability that a CPS ASEC respondent receives a PIK amounts to calculating the probability that a tax-filing CPS ASEC respondent is found in the Form 1040 data. Table A3 compares the number of filers from each year's Form 1040 file, which can be considered the population of filers, with the weighted number of filers found in the CPS ASEC after reweighting.

TABLE A3. Comparison of Form 1040 Counts and Weighted CPS-ASEC Persons Linked to Form 1040 Data, by Tax Year

Tax Year	1040 Filers	CPS ASEC weighted filers	Proportion covered
2008	138,833,790	130,203,505	0.938
2009	136,554,347	127,939,439	0.937
2010	139,057,456	127,430,267	0.916
2012	141,491,343	131,635,761	0.930
2013	142,906,142	130,800,083	0.915
2014	144,417,573	131,308,372	0.909

SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014.

Once the weights are recalculated, the number of CPS ASEC tax filers is between 91 and 94 percent of total tax filers in the 1040 file. There are several reasons why we might not cover the total number of tax filers. First, CPS ASEC respondents with imputed income data are not included in the EITC eligibility and take-up calculations. Essentially, with several variables that are key to estimation, a survey value is necessary (an example is investment income). The second issue is that Form 1040 filers may file from outside of the country. Estimates of Americans living and working abroad range from 2.2 to 6.8 million people (Costanzo and Klekowski von Koppenfels (2013)).

The statistics can be compared with an analysis of 2010 Census data (the decennial), which was linked by all persons in the Form 1040 (primary, secondary, and dependent PIKs). About 90 percent of all Form 1040 persons were found in 2010 decennial data. In the Form 1040-decennial data match, about 74 percent were White alone, about 12 percent were Black alone, and about 5 percent were Asian (which compares with 81 percent, 11 percent, and 5 percent in the present study). This higher rate for White alone may be due to primary filer characteristics differing from all persons, or due to differences in question wording between the decennial and CPS ASEC, or both. About 15 percent of persons in the Form 1040-decennial match were of Hispanic origin, compared with 14 percent in this study. Race and Hispanic origin were the only variables that were examined in the two studies.

3. Full Regression Results

Tables A4 and A5 show the full results of the main regressions reported in the paper. The only excluded variables are the State fixed effects.

TABLE A4. Results of Difference Analysis, Weighted

Dependent variable is the probability of incorrect payment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison	(4) Low income
Preparer	0.020 ***	0.011 ***	0.002	0.010 ***
	(0.003)	(0.002)	(0.003)	(0.003)
Online	0.018 ***	0.009 ***		0.012 **
	(0.003)	(0.003)		(0.004)
Preparation X Product	0.209 ***	0.134 ***	0.083 ***	0.128 ***
	(0.013)	(0.012)	(0.011)	(0.012)
Online X Product	0.078 ***	0.050 ***		0.056 ***
	(0.012)	(0.011)		(0.014)
Preparer X 2009	0.004	0.004	-0.004	0.009
	(0.005)	(0.004)	(0.004)	(0.005)
Preparer X 2010	0.003	0.000	-0.019 ***	0.003
	(0.003)	(0.004)	(0.004)	(0.006)
Preparer X 2012	0.001	-0.003	-0.006	-0.001
	(0.004)	(0.004)	(0.003)	(0.005)
Preparer X 2013	0.011 ***	0.004	-0.009 **	0.008 *
	(0.003)	(0.003)	(0.003)	(0.004)
Preparer X 2014	0.005	-0.003	-0.002	0.003
	(0.004)	(0.005)	(0.003)	(0.007)
Online X 2009	0.009*	0.008*		0.013*
	(0.004)	(0.004)		(0.005)
Online X 2010	0.023 ***	0.019 ***		0.028 ***
	(0.004)	(0.005)		(0.006)
Online X 2012	0.008	0.003		0.005
	(0.004)	(0.004)		(0.006)
Online X 2013	0.021 ***	0.014 ***		0.021 ***
	(0.004)	(0.004)		(0.006)
Online X 2014	0.009*	-0.001		0.002
	(0.004)	(0.004)		(0.007)
Preparer X Product X 2009	0.002	0.002	-0.002	-0.002
	(0.011)	(0.011)	(0.013)	(0.012)
Preparer X Product X 2010	0.060 ***	0.062 ***	0.032	0.061 ***
	(0.009)	(0.009)	(0.017)	(0.010)
Preparer X Product X 2012	-0.018*	-0.021*	-0.018	-0.022*
	(0.008)	(0.008)	(0.014)	(0.009)
Preparer X Product X 2013	-0.019	-0.018	-0.033	-0.016
	(0.013)	(0.012)	(0.017)	(0.013)
Preparer X Product X 2014	-0.002	0.001	-0.015	-0.000
	(0.012)	(0.011)	(0.016)	(0.011)
Online X Product X 2009	0.002	0.004		0.002
	(0.012)	(0.011)		(0.015)
Online X Product X 2010	0.029*	0.030*		0.030
	(0.013)	(0.013)		(0.016)

Footnotes at end of table.

Table A4. Results of Difference Analysis, Weighted—(Continued)

Dependent variable is the probability of incorrect payment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison	(4) Low income
Online X Product X 2012	0.002	-0.003		-0.008
	(0.012)	(0.012)		(0.016)
Online X Product X 2013	0.018	0.014		0.013
	(0.013)	(0.013)		(0.016)
Online X Product X 2014	0.015	0.016		0.020
	(0.012)	(0.011)		(0.014)
Product			0.050 ***	
			(0.011)	
Product X 2009			0.004	
			(0.011)	
Product X 2010			0.030 *	
			(0.013)	
Product X 2012			-0.003	
			(0.012)	
Product X 2013			0.014	
			(0.013)	
Product X 2014			0.016	
			(0.011)	
Year=2009	0.003	0.001	0.009 **	0.000
	(0.003)	(0.003)	(0.003)	(0.003)
Year=2010	0.008*	0.010 *	0.030 ***	0.013 *
	(0.003)	(0.004)	(0.004)	(0.006)
Year=2012	0.002	0.005	0.008 **	0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Year=2013	-0.004	0.003	0.016 ***	0.000
	(0.004)	(0.004)	(0.003)	(0.005)
Year=2014	0.002	0.012*	0.012 ***	0.011
	(0.005)	(0.005)	(0.002)	(0.007)
Age		0.006 ***	0.006 ***	0.009 ***
		(0.000)	(0.000)	(0.000)
Age squared		-0.000 ***	-0.000 ***	-0.000 ***
		(0.000)	(0.000)	(0.000)
Sex		-0.039 ***	-0.042 ***	-0.036 ***
		(0.004)	(0.005)	(0.004)
Head of household		-0.022 ***	-0.025 ***	-0.070 ***
		(0.006)	(0.006)	(0.006)
Married		0.023 ***	0.024 ***	0.025 ***
		(0.003)	(0.003)	(0.003)
One child		0.073 ***	0.073 ***	0.134 ***
		(0.004)	(0.004)	(0.004)
Two children		0.063 ***	0.063 ***	0.125 ***
		(0.003)	(0.002)	(0.003)

Footnotes at end of table.

TABLE A4. Results of Difference Analysis, Weighted—(Continued)

Dependent variable is the probability of incorrect payment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison	(4) Low income
Three or more children		0.042 ***	0.042 ***	0.089 ***
		(0.003)	(0.003)	(0.007)
Black		0.068 ***	0.070 ***	0.075 ***
		(0.007)	(0.007)	(0.007)
Asian		0.007 *	0.008 **	0.017 ***
		(0.003)	(0.003)	(0.004)
Other		0.009 **	0.009 *	0.011 *
		(0.004)	(0.004)	(0.005)
Hispanic		0.043 ***	0.043 ***	0.040 ***
		(0.003)	(0.003)	(0.003)
Native-born citizen		0.021 ***	0.021 ***	0.016 ***
		(0.002)	(0.002)	(0.003)
HS degree		-0.006 *	-0.007 *	-0.005
		(0.003)	(0.003)	(0.003)
Some college		-0.024 ***	-0.025 ***	-0.022 ***
		(0.003)	(0.003)	(0.003)
BA/BS+		-0.035 ***	-0.037 ***	-0.035 ***
		(0.004)	(0.004)	(0.004)
Self-employed		0.131 ***	0.122 ***	0.091 ***
		(0.015)	(0.016)	(0.017)
Income zero or less		0.067 ***	0.067 ***	0.056 ***
		(0.003)	(0.003)	(0.004)
Log income		-0.005 ***	-0.005 ***	-0.004 ***
		(0.000)	(0.000)	(0.000)
Log income squared		0.107 ***	0.109 ***	0.153 ***
		(0.005)	(0.005)	(0.007)
No TANF		0.030	0.022	0.035
		(0.092)	(0.093)	(0.093)
Log TANF value		0.005	0.004	0.006
		(0.012)	(0.012)	(0.012)
No SNAP		-0.021	-0.025	-0.022
		(0.013)	(0.014)	(0.013)
Log SNAP value		0.004	0.004	0.000
		(0.002)	(0.003)	(0.003)
Urban		-0.013 ***	-0.012 ***	-0.011 **
		(0.002)	(0.002)	(0.003)
Constant	-0.004	-0.310 **	-0.286 **	-0.326 ***
	(0.003)	(0.089)	(0.090)	(0.087)
N	336,166	336,166	313,632	211,490

* p<0.05, ** p<0.01, *** p<0.001.

SOURCE: Linked CPS ASEC-Form 1040 data, 2008-2010, 2012-2014. Results from a difference-in-differences model comparing five groups: paper filers as the base category; preparer with a product; online with a product; and preparer and online with no product. Standard errors clustered at the State level shown in parentheses.

TABLE A5. Results of Difference Analysis, GLM With a Log Link, Weighted

Dependent variable is the value of overpayment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison
Preparer	0.836 ***	0.663 ***	0.002
	(0.122)	(0.126)	(0.003)
Online	0.823 ***	0.641 ***	
	(0.131)	(0.131)	
Preparation X Product	1.861 ***	0.885 ***	0.083
	(0.082)	(0.055)	(0.011)
Online X Product	1.173 ***	0.714 ***	
	(0.128)	(0.122)	
Preparer X 2009	0.070	0.056	-0.004
	(0.162)	(0.150)	(0.004)
Preparer X 2010	0.100	0.012	-0.019
	(0.166)	(0.185)	(0.004)
Preparer X 2012	-0.179	-0.304	-0.006
	(0.225)	(0.218)	(0.003)
Preparer X 2013	0.246	0.005	-0.009
	(0.191)	(0.201)	(0.003)
Preparer X 2014	-0.036	-0.280	-0.002
	(0.284)	(0.286)	(0.003)
Online X 2009	0.097	0.097	
	(0.162)	(0.150)	
Online X 2010	0.351	0.258	
	(0.184)	(0.204)	
Online X 2012	0.066	-0.049	
	(0.228)	(0.225)	
Online X 2013	0.416 *	0.197	
	(0.198)	(0.203)	
Online X 2014	0.090	-0.153	
	(0.262)	(0.267)	
Preparer X Product X 2009	-0.041	0.007	-0.002
	(0.073)	(0.063)	(0.013)
Preparer X Product X 2010	0.050	0.117	0.032
	(0.079)	(0.072)	(0.017)
Preparer X Product X 2012	-0.094	-0.053	-0.018
	(0.089)	(0.095)	(0.014)
Preparer X Product X 2013	-0.146	-0.032	-0.033
	(0.079)	(0.066)	(0.017)
Preparer X Product X 2014	-0.075	0.047	-0.015
	(0.079)	(0.073)	(0.016)
Online X Product X 2009	-0.095	-0.068	
	(0.141)	(0.140)	
Online X Product X 2010	-0.228	-0.205	
	(0.154)	(0.138)	

Footnotes at end of table.

TABLE A5. Results of Difference Analysis, GLM With a Log Link, Weighted—(Continued)
Dependent variable is the value of overpayment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison
Online X Product X 2012	-0.129 ***	-0.174	
	(0.173)	(0.172)	
Online X Product X 2013	-0.067 ***	-0.062	
	(0.146)	(0.147)	
Online X Product X 2014	-0.094 ***	-0.073	
	(0.157)	(0.145)	
Product			0.719 ***
			(0.122)
Product X 2009			-0.070
			(0.140)
Product X 2010			-0.207
			(0.138)
Product X 2012			-0.175
			(0.172)
Product X 2013			-0.063
			(0.147)
Product X 2014			-0.075
			(0.146)
Year=2009	0.082	0.048	0.145 *
	(0.123)	(0.114)	(0.073)
Year=2010	0.159	0.232	0.492 ***
	(0.174)	(0.190)	(0.074)
Year=2012	0.317	0.414*	0.365 ***
	(0.195)	(0.194)	(0.085)
Year=2013	0.067	0.251	0.448 ***
	(0.203)	(0.215)	(0.059)
Year=2014	0.361	0.597 *	0.444 ***
	(0.276)	(0.283)	(0.086)
Age		0.167 ***	0.165 ***
		(0.007)	(0.007)
Age squared		-0.002 ***	-0.002 ***
		(0.000)	(0.000)
Sex		-0.812 ***	-0.816 ***
		(0.032)	(0.033)
Head of household		-0.928 ***	-0.935 ***
		(0.040)	(0.039)
Married		0.233 ***	0.229 ***
		(0.027)	(0.027)
One child		0.710 ***	0.694 ***
		(0.035)	(0.036)
Two children		0.780 ***	0.765 ***
		(0.041)	(0.043)

Footnotes at end of table.

TABLE A5. Results of Difference Analysis, GLM With a Log Link, Weighted—(Continued)
Dependent variable is the value of overpayment

Item	(1) Baseline	(2) With covariates	(3) Online as comparison
Three or more children		0.684 ***	0.675 ***
		(0.057)	(0.058)
Black		0.860 ***	0.854 ***
		(0.048)	(0.049)
Asian		0.365 ***	0.374 ***
		(0.038)	(0.039)
Other		0.226 ***	0.227 ***
		(0.057)	(0.057)
Hispanic		0.606 ***	0.599 ***
		(0.052)	(0.052)
Native-born citizen		0.177 ***	0.177 ***
		(0.052)	(0.052)
HS degree		-0.038	-0.038
		(0.021)	(0.021)
Some college		-0.341 ***	-0.343 ***
		(0.018)	(0.019)
BA/BS+		-0.862 ***	-0.859 ***
		(0.042)	(0.043)
Self-employed		12.485 ***	12.397 ***
		(1.617)	(1.628)
Income zero or less		3.234 ***	3.208 ***
		(0.343)	(0.348)
Log income		-0.184 ***	-0.183 ***
		(0.018)	(0.019)
Log income squared		0.950 ***	0.938 ***
		(0.036)	(0.038)
No TANF		0.057	0.022
		(0.627)	(0.619)
Log TANF value		0.024	0.020
		(0.080)	(0.079)
No SNAP		-0.153	-0.147
		(0.112)	(0.113)
Log SNAP value		-0.018	-0.018
		(0.021)	(0.022)
Urban		-0.142 ***	-0.115 ***
		(0.035)	(0.034)
Constant	2.702 ***	-13.000 ***	-12.179 ***
	(0.116)	(2.013)	(1.989)
N	336,166	336,166	313,632

* p<0.05, ** p<0.01, *** p<0.001.
SOURCE: Linked CPS ASEC-Form 1040 data, 2008–2010, 2012–2014. Results from a difference-in-differences model comparing five groups: paper filers as the base category; preparer with a product; online with a product; and preparer and online with no product. Standard errors clustered at the State level shown in parentheses.

4

Creative Use of Nontax Data Sources

Curt Hopkins ◆ Ken Su

Saurabh Datta ◆ Patrick Langetieg ◆ Brenda Schafer

Daniel Berger ◆ Eric Toder ◆ Victoria Bryant

John Guyton ◆ Patrick Langetieg

Patrick Langetieg ◆ Mark Payne ◆ Alan Plumley

Supplementing IRS Data with External Credit Report Data in Employment Tax-Predictive Models

Curt Hopkins and Ken Su (IRS, Small Business/Self-Employed Division)

Introduction

This project helped determine the value of adding business credit reporting data to existing IRS data in forecasting the likelihood of unpaid employment taxes. A credit bureau[1] provided business credit data and their three business credit scores for use in this project. We took data analysis and modeling approaches to reveal the relationship between unpaid employment tax and credit bureau credit information.

We used a specific point in time and balance due to define a noncompliant Form 941 employment tax return. Our time period target is the fourth quarter of 2012. The external data set includes eight quarters of information prior to this quarter. We define a noncompliant return as one with a balance due at first notice of at least $1,000 for data exploration and $5,000 for modeling. About 17.7 percent of records in our data set of about 288,000 businesses were noncompliant per the lower threshold and 7.3 percent per the higher threshold (unweighted).

We chose the employment tax arena for three reasons. First, prediction of employment tax liabilities is well understood by SB/SE Collection staff. Second, quarterly returns match the frequency of the credit scores in the provided dataset. Third, quarterly returns allowed for prediction across shorter time horizons, reducing the effect of external economic events.

Data Description

This project involved data from two sources: the IRS and a credit bureau. IRS data represent a sunk cost for research; it will be paid for whether these data tables are used in a specific model or not. The credit bureau data were available under an existing contract with Research, Applied Analytics, and Statistics (RAAS). Future use of credit bureau data will have an additional cost; if it benefits predictive models then a return on investment calculation will be appropriate.

IRS Data

The IRS data used in this project encompassed transcribed return data, payments, subsequent condition and transaction data, and other indicators regarding businesses the IRS already tracks in order to administer the tax system.

External Credit Bureau Data

The data from the credit bureau were new to this project and require more description. The data were provided to us by the RAAS staff based upon a stratified sample[2] devised by them for another project. The sample included 32 strata designed to represent many market subsegments including:

[1] We do not wish to name the specific credit bureau used. It was a major business in the market.

[2] The 32 strata weights were not provided, so all analyses and models are unweighted. This could have had a significant impact on the results.

- Four categories based on number of employees;
- Sole proprietorships and other business types;
- Semiweekly depositors and other deposit requirements; and
- Four definitions of compliance issues.

We used three business credit risk scores and nineteen other pieces of credit reporting data. The three credit scores include an overall business creditworthiness score (CredScore), one for financial risk prior to granting credit (FinRisk), and one to rate businesses that already owe a debt (CollPred). None of these is a product name from the credit bureau.

The first two scores are scaled from 1 to 100, with lower scores representing poorer credit.[3] The CredScore looks at risk in terms of severe payment delinquency. The credit bureau optimizes FinRisk to predict the potential to default on an obligation. CollPred ranges from 100 to 900, with the lower scores correlated with the lowest likelihood of payment. However, CollPred is not broken into classes (see below); it is purely an ordinal score.

The credit bureau summarizes the first two scoring systems into Risk Classes. Those Risk Classes with the lowest scoring businesses include a higher percentage of poor-performing businesses as defined by the target of the models. CredScore poor performers would have severe payment delinquency and FinRisk poor performers would be those defaulting on a debt. Table 1 summarizes the Risk Classes.

TABLE 1. Risk Classes and Percent of Poor-Performing Businesses

Risk Class	CredScore	Delinquents	FinRisk Score	Defaulters
High	1–10	50.8%	1–3	35.3%
High-Medium	11–25	19.1%	4–10	10.0%
Medium	26–50	10.0%	11–30	2.9%
Low-Medium	51–75	4.4%	31–65	1.1%
Low	76–100	1.7%	66–100	0.6%

SOURCE: Credit bureau white papers

We include 19 key business credit risk factors from the external data set in our analysis:

1. Count of new trades under delinquency;
2. Count of continuous trades under delinquency;
3. Days Beyond Terms (DBT) of combined trades;
4. Count of aged trades;
5. Count of aged trades under delinquency;
6. Count of trades under Days Beyond Terms (DBT);
7. Count of total trades under delinquency;
8. Count of banking liability relationship;
9. Count of leasing trades;
10. Count of leasing trades under delinquency;
11. DBT of additional trades reported within the last 4 months;
12. Count of UCC[4] filings—reported as write-offs or skips;

[3] There are also scores indicating missing and out-of-range values. We treated those scores as missing data in this project.

[4] Uniform Commercial Code filings, required whenever a company pledges assets as collateral.

13. Count of unsatisfied UCC filings within the last 24 months;
14. Count of legal filings in past 6 months;
15. Count of bankruptcies filed within 6 months;
16. Count of tax liens filed within 6 months;
17. Count of judgments filed within 6 months;
18. Count of open and closed collection trades placed within 12 months; and
19. Total number of inquiries in the last 3 full months.

Findings

We present our findings in three types of detailed analyses:

1. In data exploration, we report bivariate tables of external data across compliant and noncompliant taxpayers.
2. Our modeling analysis includes three phases: modeling from IRS data; modeling with both IRS and credit bureau data; and predicting changes in credit bureau credit risk using IRS data.
3. The cause and effect analysis tests if there is a causal link between the credit reporting and IRS data as a time series.

1. Credit Data Exploration

In this section, we explore the relationship between the credit bureau data and employment tax compliance. The data include the nineteen credit risk factors and three business credit scores described above. We prepared the data for analysis as follows:

A. We determined the compliance level of each business in each quarter.
   i. Noncompliant businesses are those with a balance due of over $1,000 at first notice.
   ii. Compliant businesses did not have a balance due (and thus no first notice).
   iii. Those with a balance due between $1 and $1,000 are not shown in this section in order to maximize the contrast between our two groups of interest.

B. We then developed an overall profile of each group's risk factors from the credit bureau business data perspective to uncover the relationships between noncompliant taxpayers and the credit bureau's data.

Table 2 demonstrates our profile of noncompliant and compliant businesses during the four quarters of 2012. The percentage in each cell quantifies the businesses meeting the credit risk factor. For example, in the very first data cell we show that in the first quarter of 2012, 0.90 percent of businesses with a balance of at least $1,000 at first notice also have at least one new delinquent trade reported by the credit bureau. This compares to the cell below, where 0.96 percent of compliant businesses had a new delinquent trade.

These results demonstrate that these credit risk indicators did *not* significantly differentiate taxpayers with a balance due from the compliant group. We confirmed this with z-tests on the larger differences. Despite these initially negative results, we allowed consideration of the credit risk indicators in our analysis by modeling.

Continuing our data exploration, we focused on the credit bureau credit scores. Three tables below show the percent of cases with a balance due in each credit bureau Risk Class for each type of credit score. For brevity purposes, the tables show the three scores for 2012, but the results are the same in 2011 and 2013. Further, the same pattern holds for balances of at least $5,000 in each year; the rate in each cell is lower, but the near-constant rate across score ranges is the same. Ranges and descriptions are those defined by the credit bureau.

TABLE 2. Percent With a Credit Risk Factor by Employment Tax Compliance Category

Compliance Category	1Q2012	2Q2012	3Q2012	4Q2012	Average
Business With a New Delinquent Trade					
Balance Due > $1,000	0.90%	0.83%	0.67%	0.53%	0.73%
Compliant	0.96%	0.88%	0.64%	0.54%	0.75%
At Least One Delinquent Continuous Trade					
Balance Due > $1,000	19.20%	19.75%	20.11%	21.32%	20.10%
Compliant	19.28%	19.97%	20.36%	21.69%	20.32%
At Least One Delinquent Trade >30 Days Beyond Terms					
Balance Due > $1,000	16.15%	16.58%	17.01%	18.02%	16.94%
Compliant	16.12%	16.60%	17.09%	18.01%	16.96%
At Least One Aged Trade					
Balance Due > $1,000	45.35%	45.55%	42.71%	43.20%	44.20%
Compliant	45.03%	44.94%	42.28%	43.02%	43.82%
At Least One Delinquent Aged Trade					
Balance Due > $1,000	12.14%	12.06%	12.97%	13.61%	12.70%
Compliant	12.36%	12.10%	12.99%	13.64%	12.77%
At Least One Trade Not Beyond Term					
Balance Due > $1,000	26.86%	27.27%	28.18%	29.58%	27.97%
Compliant	26.71%	27.21%	28.10%	29.58%	27.90%
Any Delinquent Trade					
Balance Due > $1,000	26.86%	27.27%	28.18%	29.58%	27.97%
Compliant	26.71%	27.21%	28.10%	29.58%	27.90%
Banking Liability					
Balance Due > $1,000	0.32%	0.34%	0.35%	0.36%	0.34%
Compliant	0.38%	0.40%	0.40%	0.41%	0.40%
Reported Leasing Trade					
Balance Due > $1,000	2.24%	2.26%	2.31%	2.41%	2.31%
Compliant	2.30%	2.33%	2.36%	2.45%	2.36%
Delinquent Leasing Trade					
Balance Due > $1,000	0.003%	0.003%	0.003%	0.005%	0.004%
Compliant	0.000%	0.000%	0.000%	0.004%	0.001%
Combined Trades under DBT					
Balance Due > $1,000				4.68%	
Compliant				4.68%	
At Least One UCC Filing Within the Last 24 Months					
Balance Due > $1,000	20.93%	21.64%	22.55%	23.03%	22.04%
Compliant	20.68%	21.35%	22.24%	22.65%	21.73%
At Least One Legal Filing in Past 6 Months					
Balance Due > $1,000	3.15%	3.48%	3.64%	3.82%	3.52%
Compliant	3.06%	3.49%	3.71%	3.83%	3.52%
Bankruptcy Filing in Past 12 Months					
Balance Due > $1,000	0.12%	0.13%	0.15%	0.17%	0.14%
Compliant	0.15%	0.15%	0.15%	0.18%	0.16%

SOURCE: IRS Compliance Data Warehouse (CDW) and credit bureau data.

TABLE 3. Unpaid Tax Rate Within Each FinRisk Risk Class

FinRisk		Percent With an Unpaid Balance > $1,000				
Score Range	Risk Class	1Q2012	2Q2012	3Q2012	4Q2012	Average
1 – 3	High	28.4%	28.3%	28.2%	27.9%	28.2%
4 – 10	High-Medium	28.9%	28.8%	28.9%	28.9%	28.9%
11 – 30	Medium	28.9%	29.2%	29.2%	29.2%	29.1%
31 – 65	Low-Medium	29.0%	28.9%	28.8%	28.7%	28.8%
66 – 100	Low	28.8%	28.8%	28.9%	29.0%	28.9%

SOURCE: IRS CDW and credit bureau data.

TABLE 4. Unpaid Tax Rate Within Each CredScore Risk Class

CredScore		Percent With an Unpaid Balance > $1,000				
Score Range	Risk Class	1Q2012	2Q2012	3Q2012	4Q2012	Average
1 – 10	High	28.7%	28.7%	28.9%	28.7%	28.8%
11 – 25	High-Medium	29.2%	29.1%	28.6%	28.8%	28.9%
26 – 50	Medium	28.9%	28.9%	29.1%	29.2%	29.0%
51 – 75	Low-Medium	28.8%	29.0%	29.0%	28.9%	28.9%
76 – 100	Low	28.9%	28.7%	28.8%	29.0%	28.9%

SOURCE: IRS CDW and credit bureau data.

The credit bureau does not provide category definitions for the CollPred Score. We created simple categories to see if the same pattern held. In this case, the percentages show those who paid a balance due within the next six months (to parallel the definition of this score provided by the credit bureau).

TABLE 5. Unpaid Tax Rate Within Each CollPred Risk Class

CollPred		Percent Paying the Balance Due Within 6 Months				
Score Range	Risk Class	1Q2012	2Q2012	3Q2012	4Q2012	Average
1–10	Very Low	6.8%	7.2%	7.3%	7.0%	7.1%
11–15	Low	7.4%	7.4%	7.5%	7.6%	7.5%
16–20	Low-Medium	7.6%	7.6%	7.5%	7.5%	7.6%
21–25	Medium	7.5%	7.5%	7.5%	7.6%	7.5%
26–50	Medium-High	7.4%	7.4%	7.4%	7.4%	7.4%
50+	High	7.7%	7.8%	7.8%	7.5%	7.7%

SOURCE: IRS CDW and credit bureau data

We found no indication that noncompliance increased with the Risk Classes defined by the credit bureau or the priority ranges set in a similar fashion. We confirmed this with chi-square tests showing that these percentages mimic a uniform distribution.

2. Analysis by Modeling

In the second set of analyses, we built models to determine the additional benefit of including credit reporting data in predicting future employment tax delinquencies.

Phase I

Initially, we built a model to predict which Forms 941 for the fourth quarter of 2012 would owe at least $5,000 at first notice using IRS information available from prior returns and other information known at the end of the prior quarter (third quarter 2012). We do include the Form 941 tax return for the third quarter of 2012, acknowledging that it is filed one month into the fourth quarter.

Phase II

Starting with the Phase I model, we then allowed consideration of credit bureau data up to the third quarter of 2012. The data included both the credit risk indicators and the three credit bureau scores for each business (detailed previously), as well as derived information such as score ranges and changes in scores across quarters.

Phase III

We then built models to predict the change in credit bureau Risk Classes from the third to fourth quarters of 2012. If the credit bureau scores predict taxpayer behavior, then we believe the relationship will hold true in the other direction and IRS data can support adequate predictive models.

General Modeling Methodology

After data preparation, we built logistic regression models using a modified stepwise method. The initial variables under consideration came from factor analysis of the available variables against the target variable. We allowed as many as 50 factors[5] in order to provide a broad selection of variables. We did not use the factors themselves, but rather selected the variable most correlated with the dependent variable from within each factor. Our intention was to have many variables available for stepwise consideration while minimizing autocorrelation among the available variables.

Methodology

We started with available information from each source, then transformed, binned, and made indicators from it, and then used a standard methodology to provide modeling variables and evaluate the results. Over one-hundred models were created at different times in this project; only results from the models determined best in their specific phase (based on diagnostic tests) are included in this report.

Variable Creation

Beginning with variables transcribed from tax returns, business entity information, and subsequent transaction and status changes, we expanded the variables by various types of recoding to support our modeling efforts. We used the following general techniques on the IRS data:

- Data Transformations

  For dollar amounts and counts of events (e.g., tax deposits, number of returns filed), we kept the raw data and added transformations by natural logs and square roots to provide three versions of each amount. We made a fourth version of tax return data by dividing dollar line items by the total wages reported on the return; this gives a less volatile amount, generally between 1 and 100 percent of the total wages.

- Data Binning

  Based on our experience with modeling payment compliance, we also converted dollar amounts to bins (ranges). This was especially useful for accounts receivable in the prior four quarters, as prior noncompliant behavior (e.g., owing $3,000 to $4,000 in the third prior quarter) is a good indicator of future noncompliance.

- Indicators

  We also set up indicators for specific conditions (e.g., prior installment agreements, prior notices of Federal tax lien, bad checks, and bankruptcy). Many of the "raw" variables from the IRS Compliance Data Warehouse (CDW) are themselves indicators (e.g., filing requirements, and specific transaction codes).

[5] No factor was included if its eigenvalue was below 1.0.

- Differences Across Quarters

  The techniques described above generally reshaped data within a specific quarter. We also compared quarters to each other and computed differences in dollar amounts, delinquencies, and counts between quarters. We set indicators for compliance events, such as a sudden drop in tax deposits or skipped filing. Finally, we set indicators and continuous variables based on the amount of variation in monthly and quarterly wages, deposits, and other payments.

Among these different techniques, we created a rich data set with over 550 variables considered for inclusion in the various models.

Credit Bureau Data

Using the ordinal ranges described above for the three credit bureau scores, we transformed this data in a similar way.[6] Raw scores were transformed by natural logarithms and square roots. We added binning of the first two scores based on the Risk Classes from the credit bureau and also created comparisons across quarters for changes—increases and drops in score ranges. In total, we added over 150 variables to the data set for the three scores from 11 quarters in our time window.

Variable Reduction

The data sets now contain many closely correlated variables and we set out to reduce these to a vital few for the models. For this, we used an existing program created by the Strategic Analysis and Modeling Group within Collection that creates factors and calculates the correlation of each member of the factor with the target variable.[7] With the factors made, we selected the variable with the greatest correlation to the target as our initial candidate and tried different numbers of factors to increase the potential candidates or reduce multicollinearity among the candidates.

Variable Selection

Variables selected for inclusion in the model were determined in iterations composed of two steps. The first step was stepwise logistic regression with an entry value of .01 and exit of .10. These values allow variables with strong predictive power into the model and tend to keep the variable in the model unless other (later) additions make it redundant. The second step required manual intervention. We evaluated the variables included in the initial model and, based on experience and graphical analysis, tested substitute variables. If the substitution resulted in better diagnostics, the change was incorporated into the next iteration of the model. It is quite possible (and did happen) that a square root transformation of a dollar variable was most closely correlated in the factor analysis, but with the inclusion of many other variables in the model, only an indicator was needed and not the continuous variable. We repeated these steps dozens of times to achieve our best models of future noncompliance and used this iterative process in each of the three modeling phases of this project.

Model Evaluation

In all models, we used nine standard (and one custom) diagnostics to determine if the model improved over the prior version. Here is a short summary of the diagnostics used in this project:

Akaike Information Criterion (AIC)

A measure of relative quality of model fit providing a means to compare models with differing numbers of independent variables.

Schwartz Criteria (SC)

SC is another model fit statistic similar to AIC, but it penalizes models more for including additional variables without a corresponding increase in predictive ability.

[6] While not strictly appropriate for ordinal data, we wanted to explore these alternatives due to the poor relationship found between the raw scores and compliance levels in the first section of our analysis.

[7] Balance due of $5,000 or more at first notice.

Somers' D

This test of rank-order (ordinal) correlation varies from -1 to +1 with zero indicating no correlation.

Area Under Curve (AUC)

The AUC measurement can be used to compare multiple models. Based on the Receiver Operator Curve, the greater this value (on a 0.5 to 1.0 scale), the better the model's predictive power. A value of 0.5 is considered to represent a random model, no better than flipping a coin.

Hosmer-Lemeshow Test (H-L)

The H-L Test compares the number of correct predictions in each decile (based on model score) to a theoretical distribution using a chi-square goodness of fit test.

Deviance

A measure used to help determine if a more complex model should be used. p-Values above 0.05 indicate that we should not reject the current model in favor of a more complex one.

Model Results Test—Percent in Top Decile

In this simulation of workload selection, the highest scores would have been sent forward for fieldwork. We compared the percentage of target cases found in the top decile as a proxy for moving the best selection (most true positives and fewest false positives) to the field.

Also considered during model development:

Global Likelihood Ratio (GLR)

This measure shows whether at least one of the independent variable coefficients is significantly different from zero.

Variance Inflation Factor (VIF)

VIF is a standard measure of multicollinearity among the independent variables in a model.

Model Development Findings

We present our modeling findings as a single table at the start of this section for ease of reference. Some findings and all conclusions draw across phases.

TABLE 6. Combined Model Performance and Diagnostics

Model	AIC	SC	Somers' D	AUC	H-L	Deviance	Top Decile Percent
Phase I: Predict Balance Due on Form 941 with IRS Data							
IRS Data Only	101,618	102,353	0.72	0.86	269.0	0.36	56.5%
Phase II: Add Credit Bureau Data to Phase I Model							
Mixed Data	101,608	102,385	0.72	0.86	265.0	0.37	56.5%
Phase IIIa: Predict Worsening Credit Bureau Risk Class with IRS Data							
FinRisk IRS Data Only	120,665	120,675	0.00	0.50	N/A	0.53	10.4%
CredScore IRS Data Only	182,746	182,787	0.01	0.51	2.9	0.80	9.9%
Phase IIIb: Predict Worsening Credit Bureau Risk Class with IRS Data and Prior Risk Class							
FinRisk Mixed Data	113,706	113,840	0.34	0.67	2.9	0.50	22.8%
CredScore Mixed Data	172,586	172,752	0.32	0.66	983.0	0.77	17.3%

Phase I. Model Based Upon IRS Data Exclusively

The model from the IRS data shows solid diagnostics. Somers' D is over 0.70 and AUC above 0.85. The Deviance is close to 0.5 (optimal) and its p-value does not indicate the need for a more complex model. Finally, over 55 percent of target cases are found in the top decile.

TABLE 7. Combined Model Performance and Diagnostics

Model	AIC	SC	Somers' D	AUC	H-L	Deviance	Top Decile Percent
Phase I: Predict Balance Due on Form 941 with IRS Data							
IRS Data Only	101,618	102,353	0.72	0.86	269.0	0.36	56.5%

Phase II. Comparing Models Without and With Credit Bureau Data

In Phase II, we take the model from Phase I and supplement those models with the credit bureau information. After initial attempts at restricted variable choices,[8] using the same factor technique described for Phase I failed to improve the existing IRS data model, we allowed all credit bureau data (raw, transformed, categorized, and indicators for changes across quarters) into consideration for the model. Among these 22 variables (and dozens of derived indicators), the stepwise selection chose four to be included in the model: one based on the most risky categories of CredScore and three indicators for changes across quarters in the FinRisk Scores. No variables from CollPred made it into the final model. The variables chosen were:

CredScore_Bad032012	This indicator is set to 1 if the business was in either of the two riskiest classes based on the first quarter 2012 CredScore.
FinCat062012_Worse	This indicator is set to 1 when the second quarter 2012 FinRisk Class is at least one level more risky than it was for the same business in the first quarter of 2012.
FinCat092012_Worse2	This indicator is set to 1 when the third quarter 2012 FinRisk Class is at least *two* levels more risky than it was for the same business in the second quarter of 2012.
FinCat092012_Better	This indicator is set to 1 when the third quarter 2012 FinRisk Class is at least one level *less* risky than it was for the same business in the second quarter of 2012. Note: This indicator has a negative coefficient in the model, indicating that this condition links to a business that is less likely to owe in their fourth quarter return.

The first three variables selected have coefficients that indicate a greater likelihood of a balance due if the business was considered a poor risk nine months earlier (first condition), or had a worsening financial rating three or six months prior (second and third conditions). The last variable shows a negative coefficient, consistent with the lower risk of a balance due at the same time the credit bureau coded an improving financial rating for the business (in the prior three months).

Knowing which indicators were included in the model, we then evaluated the additional predictive power and stability of the model with this new information.

TABLE 8. Combined Model Performance and Diagnostics

Model	AIC	SC	Somers' D	AUC	H-L	Deviance	Top Decile Percent
Phase I: Predict Balance Due on Form 941 with IRS Data							
IRS Data Only	101,618	102,353	0.72	0.86	269.0	0.36	56.5%
Phase II: Add Credit Bureau Data to Phase I Model							
Mixed Data	101,608	102,385	0.72	0.86	265.0	0.37	56.5%

[8] That is, choices restricted to a single variable within each of 30 factors made from these credit bureau variables.

Akaike Information Criterion (AIC)

The addition of the credit bureau data improved the AIC by 10 points—dropping from 101,618 to 101,608.

Schwartz Criteria (SC)

As a modification of the AIC that penalizes models for including additional variables without a corresponding increase in predictive ability, the SC rose (i.e., worsened) by 32 points with the addition of the credit bureau data. The extra four variables moved the rating from 102,353 to 102,385.

Somers' D

This test of rank-order (ordinal) correlation varies from -1 to +1 with zero indicating no correlation. Both the IRS and IRS with credit bureau data models have a Somers' D of 0.72 indicating very good ordinal correlation.

Area Under Curve (AUC)

The addition of the credit bureau variables to the final IRS model did not change the AUC rating. It is 0.86 under both models.

Hosmer-Lemeshow Test (H-L)

Neither model had an H-L Test that would be considered acceptable. The H-L test is satisfied when the models do not reject the assumption that their decile distribution is identical to a theoretical one when tested using chi-square. In this test, both models are too good at placing the desired cases in the top decile and keeping them out of the bottom one. The H-L theoretical top decile is assumed to have 10,486 cases and the lowest 261. The models each place over 11,110 cases in the top decile and fewer than 180 in the bottom decile—with similar results in the middle eight deciles. In other words, the models performed so well at separating the desired cases by score that the chi-square test shows independence.

Deviance

This is a measure to help determine if a more complex model should be used. P-Values above 0.05 indicate that we should not reject the current model in favor of a more complex one. The IRS model has a Deviance 0.36 and the Mixed Data model of 0.37. Both p-values approach 1.0. The models are sufficiently complex.

Model Results Test—Percent in Top Decile

In this simulation of workload selection, the highest scores would have been sent forward for fieldwork. The IRS model placed 56.5 percent of the target cases (11,118 of 19,680) in the top (highest-scoring) decile. The Mixed Data model improved this slightly, placing 11 additional cases in the top decile (still 56.5 percent).

Global Likelihood Ratio (GLR)

This measure shows whether at least one of the independent variable coefficients is significantly different from zero. In comparing models, a higher score can be interpreted as having greater significance. The IRS model's GLR is 39,409 and the Mixed Data model is 39,426. The p-value of each is <0.0001, indicating that at least some of the variables in each model are significant predictors. In fact, the Wald statistics for each predictor in the IRS Data model have a p-value below 0.029, and in the Mixed Data model below 0.035, with each predictor rejecting the hypothesis that it adds no value to the model.

Variance Inflation Factor (VIF)

The addition of credit bureau data did not change the highest VIF among variables in the model. In the Mixed Data model, the same two variables[9] had the highest VIFs, although the order of them swapped. The ratings stayed on either side of 6.0, indicating a small concern with multicollinearity. This rating is not shown in the table.

[9] The two variables are indicators that: 1) the business owed at least $4,000 more in the first prior quarter than in the third; and 2) it owed at least $6,000 in the second prior quarter.

Phase III. Predicting Credit Bureau Scores or Categories Using IRS Data

If the credit bureau data were closely linked to tax compliance, it should be possible to predict certain features of the credit bureau scores using IRS tax data. Under this assumption, we tried to model the condition that the credit bureau score was in a worse Risk Class in the fourth quarter of 2012 than it was in the third. We chose movement between Risk Classes since this information was selected into the models in Phase II and should, therefore, be the most similar to IRS tax information. Only two of the three credit scores have Risk Classes.

TABLE 9. Combined Model Performance and Diagnostics

Model	AIC	SC	Somers' D	AUC	H-L	Deviance	Top Decile Percent
Phase IIIa: Predict Worsening Credit Bureau Risk Class with IRS Data							
FinRisk IRS Data Only	120,665	120,675	0.00	0.50	N/A	0.53	10.4%
CredScore IRS Data Only	182,746	182,787	0.01	0.51	2.9	0.80	9.9%
Phase IIIb: Predict Worsening Credit Bureau Risk Class with IRS Data and Prior Risk Class							
FinRisk Mixed Data	113,706	113,840	0.34	0.67	2.9	0.50	22.8%
CredScore Mixed Data	172,586	172,752	0.32	0.66	983.0	0.77	17.3%

Using just IRS data, the initial models (labeled IRS Data Only) were disappointing. Allowing one piece of credit bureau data into the mix made a large difference.

The first attempt at a FinRisk model was intercept-only (no independent variables selected). Only the intercept came in through stepwise regression, even when the level for entry eased to 0.10 and the stay criterion eased to 0.20. This held true even with all 556 IRS variables allowed into consideration. The AUC is indistinguishable from random chance and the other diagnostics are not much better. The initial CredScore model had a similar, though less extreme, set of diagnostics. Results with Somers' D at 0.01 and AUC at 0.51 are likewise not acceptable.

In the second part of this Phase, we allowed the Risk Class from the third quarter of 2012 into each model. We tried this as both ordinal and continuous data with very similar results. The continuous version is presented here. Providing the model with the prior Risk Class and predicting a more risky category the next quarter was more successful, providing a weak model. The overall model criteria (AIC and SC) came down, indicating better fit. The prediction diagnostics (Somers' D and AUC) moved to levels considered to indicate a poor model. It is important to note that almost all the improvement came from adding a piece of information not part of the IRS' tax data – the prior credit bureau Risk Class.

3. Granger Causality Testing

We built new unpaid employment tax and related credit bureau credit score time series data to perform cause and effect testing. The Granger Causality Test is a statistical hypothesis test to determine whether one time series is useful in forecasting another. The results of the Granger Causality Test are shown in the tables below.

TABLE 10. Granger—Causality Wald Test (Direction: IRS Data Predict Credit Bureau Data)

Using These Data	To Predict These Data	Chi-Square	Pr > ChiSq
IRS Compliance	CredScore	0.96	0.33
IRS Compliance	FinRisk	0.01	0.91
IRS Compliance	CollPred	0.24	0.63

TABLE 11. Granger-Causality Wald Test (Direction: Credit Bureau Data Predict IRS Data)

Using These Data	To Predict These Data	Chi-Square	Pr > ChiSq
CredScore	IRS Compliance	0.60	0.44
FinRisk	IRS Compliance	2.17	0.14
CollPred	IRS Compliance	0.72	0.40

The Wald tests show no evidence to reject the hypothesis that IRS compliance data and credit bureau scores are independent of each other at the 0.05 significance level. None of these three credit score time series can forecast our noncompliant taxpayer group or the other way around. Our results demonstrate that both data series are highly orthogonal. Credit bureau business data appear to have little predictive power in forecasting future employment tax noncompliance.

Conclusions

Our initial data exploration of credit bureau credit risk indicators and credit score risk categories showed little evidence that these data related to employment tax compliance. The inclusion of credit bureau data added a very small amount to the predictive power of the IRS-only model, resulting in just 0.1 percent more target cases in the top decile (Phase II). The six diagnostic tests are split, with one (AIC) showing an improvement, four showing no discernable change, and the last (SC) indicating that the additional predictive power does not justify the inclusion of four additional variables. Tests in the reverse direction show that IRS tax administration data and credit bureau scores show little, if any, relationship (Phase III).

Our final analysis showed no ability for credit bureau data to predict employment tax compliance, nor could IRS compliance data predict credit bureau credit scores. While not truly orthogonal, it appears that there is little relation between credit bureau business credit risk indicators or credit scores and the likelihood of owing employment taxes in the future.

Recommendations for Further Research

There is little evidence that including credit bureau data would lead to improved models for predicting employment tax compliance. Therefore, incorporating credit bureau data appears worthwhile only if there is no cost (either direct or opportunity) in bringing that data into the IRS for use in employment tax prediction. Because there is a measurable benefit in one area, it may be worthwhile to explore modeling outside the employment tax arena or testing individual (nonbusiness) credit scores at some point in the future.

Better Identification of Potential Employment Tax Noncompliance Using Credit Bureau Data

Saurabh Datta, Patrick Langetieg, and Brenda Schafer (IRS Research, Applied Analytics and Statistics) [1]

The U.S. income tax system generally requires employers to deposit payroll taxes using either a monthly or semiweekly schedule, depending on the total payroll tax liability the employer reported during the relevant look back period.[2] Any employer that accumulates $100,000 or more in payroll taxes on any day during its monthly or semiweekly deposit period must deposit the tax by the next business day. A failure to make timely deposits results in a failure-to-deposit penalty up to 15 percent and the employer may be subject to subsequent enforcement actions.

Because the IRS does not know the date on which a business makes payroll, the date of the businesses employment tax return filing sets the timeline for the expected next filing. If a return is not subsequently secured, the business is considered potentially noncompliant. This method of identification is not timely, and it is certainly not proactive. In the coming years, the Electronic Federal Tax Payment System (EFTPS) will be used to analyze the past deposits to identify deviations in the deposit patterns to determine potential noncompliance. Until then, the IRS should explore other early detection methods.

If the IRS can more quickly identify employers that have missed—or may miss—one or more scheduled payroll deposits before these employers encounter further financial difficulties, file for bankruptcy, or close down the business, then there is a higher likelihood of collecting the unpaid taxes. This early intervention may even provide useful guidance to the employer regarding the importance of payroll tax compliance and the availability of payment plans that can help the employer meet their payroll deposit responsibilities.

Previous studies have attempted to identify and examine factors that drive noncompliance in the form of tax delinquency and untimely tax deposits in the payroll tax compliance program. These efforts have been limited to IRS internal administrative data and a few surveys (SB/SE (2011); Hopkins and Su (2017)). Hopkins and Su (2017) concludes that including credit bureau data with the IRS administrative data does not contribute to the predictive accuracy of future noncompliance.

This paper discusses preliminary findings that show that matching a homogenous sample of employers with third-party short- and long-term credit bureau credit scores may be useful to proactively identify potential noncompliant employers.

The following sections highlight the sample design and data description for this study, and then the interaction of short- and long-term credit scores with employer payroll tax compliance. The final section summarizes the preliminary findings and discusses opportunities for further research and extensions of this study.

Sample Design

In a previous study undertaken by Hopkins and Su (2017), a sample of 300,000 employers was drawn to analyze the link between detection of potential payroll tax noncompliance using credit bureau data. The sampled employers varied widely in terms of their business capitalization and state of noncompliance.

---

[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] IRS Publication 15, Employer's Tax Guide (Circular E), p. 25, available at https://www.irs.gov/pub/irs-prior/p15--2017.pdf. For Form 943 filers, see IRS Publication 51, Circular A, Agricultural Employer's Tax Guide, p. 15, available at https://www.irs.gov/pub/irs-pdf/p51.pdf.

A stratified sample was drawn and sent to the credit bureau. The sample segmented businesses based on the number of employees (0 to 10, 11 to 25, 26 to 50, and over 50). They were further segmented by total tax liability (less than $2,000, $2,001 to $6,000, $6,001 to $20,000, and over $20,000). Noncompliance was defined based on two criteria: 1) assessment of a failure to deposit penalty; and 2) an unpaid payroll tax liability after receipt of the first IRS notice. The sampling criteria were further restricted to ensure that the sample was statistically representative of all the required variables over 20 quarters starting from the third quarter of Calendar Year 2010 (2010Q3). Unfortunately, the two noncompliance measures intended for this study turned out to be based on rare events in the drawn sample. Even after adequately representing businesses by number of employees and tax liabilities, assessment of a failure to deposit penalty was detected in just four percent of the sample, and an unpaid tax liability after receipt of first IRS notice was realized in only two percent of the sample. The limited occurrence of predicted noncompliance severely restricted the analysis of underlying issues driving noncompliance.

The limitations of the previous sample provided motivation to draw another random sample of 250,000 employers. The reference point for this study is the fourth quarter of Calendar Year 2014 (2014Q4). The study period consists of the eight quarters immediately before and the eight quarters immediately following 2014Q4.

The study sample was comprised of 70 percent "detected noncompliant employers." A detected noncompliant employer is defined as an employer who received a first notice regarding potentially unpaid payroll taxes at some point during the eight quarters prior to 2014Q4 and whose case ultimately resolved in an assessment of unpaid payroll taxes. The detected noncompliant cases were oversampled to allow us to study this population in greater detail. Note that the remaining 30 percent of the sample is not necessarily payroll tax compliant. They simply were not subject to enforcement action during the eight quarters prior to 2014Q4. The sample was restricted to small businesses with assets below $10 million that filed Form 941. In addition, businesses in the sample must have existed prior to January 1, 2013, to ensure that they had some credit history prior to the study reference date.

The credit bureau matched 160,627 of the 250,000 sampled employers with credit data. Thus, the final sample is comprised of 67 percent detected noncompliant employers and 33 percent other employers.

Data Description

The credit bureau has extensive data, including short- and long-term credit scores, and other credit-risk variables, such as total outstanding balance, Federal and State liens, number of outstanding legal issues, information about credit accounts in collection status, etc. Additionally, the database provides a large collection of firmographic variables, such as industry, business size, location, etc. This study concentrates specifically on the short- and long-term credit scores. The additional variables will be used as an extension to this research.

The short-term credit score predicts the likelihood of defaulting in the next 12 months on a credit obligation that has been past due for more than 91 days. This score makes use of business and consumer variables such as payment history, frequency of payments, and short-term delinquent balances. The score is computed on a scale of 1 to 100, where a higher score is associated with lower risk.

The long-term credit score predicts the probability of bankruptcy or the prospect of defaulting on 75 percent of the credit obligations that are more than 91 days past due. The score is computed using trade, public records, and firmographic data. The primary factors affecting this score are high utilization of credit lines, tax liens and judgments, and bankruptcy filings. As with the short-term credit score, the scale is 1 to 100 with a higher score meaning lower risk.

Industry lenders use both the short- and long-term credit scores when making lending decisions, determining interest rates and risk policies for the businesses (Experian (2016a and 2016b)). The collective use of both the scores provides lenders with important details about the current status of a business and the risk involved in its operations, both in the short- and long-run. The following matrix summarizes the risk categories based on the application of the two credit scores:

TABLE 1. Risk Classification Matrix

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment	Medium Risk
High	Slow Recovery	High Risk

SOURCE: RAAS Taxpayer Behavior Lab, May 2017; Experian (2016b).

In Table 1, each category has distinct significance in assessing risk and its potential future implications. Businesses that fall into the Low Short-Term Risk/Low Long-Term Risk category are stable businesses. The businesses in the Low Short-Term Risk/High Long-Term category are able to fulfill their short-term credit obligations but are falling behind in meeting their long-term credit payments. These businesses are considered medium risk. High Short-Term Risk/Low Long-Term Risk businesses are experiencing difficulties in keeping up with their short-term credit obligations but have not been defaulting on long-term credit responsibilities. These businesses are using payment plans to pay off short-term debt, thus they are considered to be in the slow recovery category (Experian (2016b)). Businesses in the fourth category, the High Short-Term Risk/High Long-Term Risk group, are in peril of financial catastrophe.

Following financial industry standards, this paper considers both the short- and long-term credit scores in its analysis. We first explore the separate relationship of the short- and long-term credit scores on the two study segments (detected noncompliant employers and other employers). We then focus the analysis on the interaction of the two credit scores with detected noncompliant employers. In the next section, the paper reports preliminary results from this analysis.

Exploratory Analysis

The credit scores from the credit bureau are classified into five risk categories (low, low-medium, medium, medium-high, and high). For simplification and comparison purposes, the paper clusters the risk categories into two broad categories (low and high) and combines the low and low-medium categories into a single low/low category (see Table 1).

In Table 2, we profile short-term credit score with detected noncompliant employers (referred to as the Detected group) and other employers for three periods: observation period (2014Q4), one year prior to the observation period (2013Q4), and two years prior to the observation period (2012Q4). The table suggests that there is no direct association between a short-term high-risk credit score and payroll tax noncompliance. The probability of being in high versus low doesn't appear to be correlated with Detected cases.

TABLE 2. Distribution of Low and High Short-Term Credit Risk Over Time

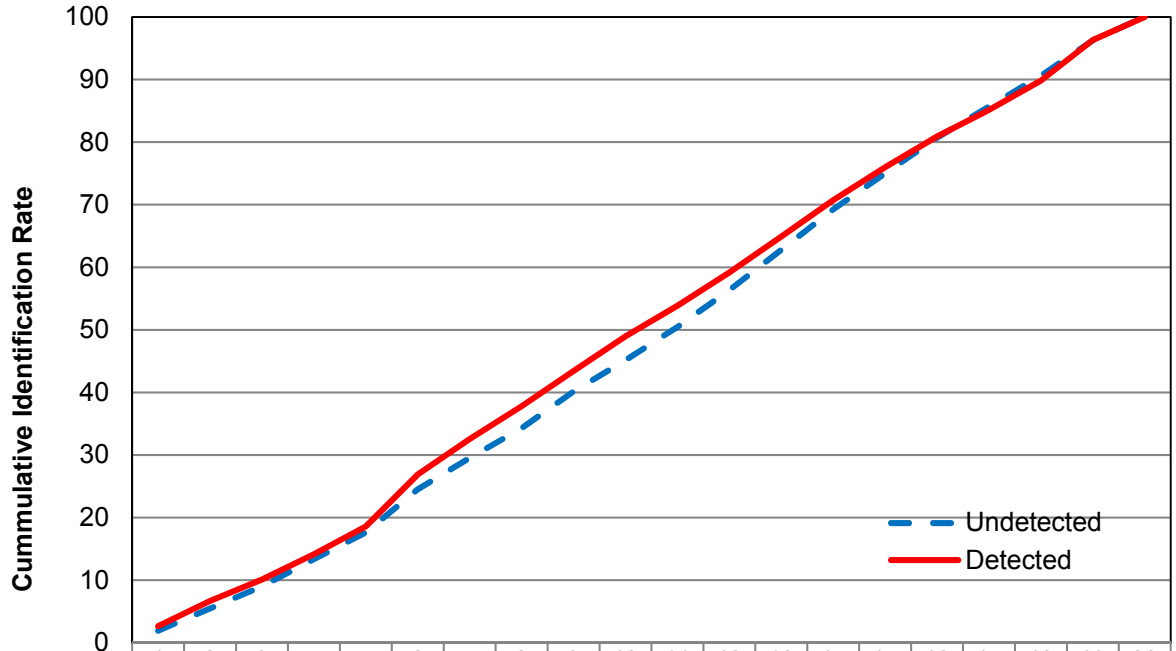
Item	2014Q4 (Observation Period)		2013Q4 (One Year Prior)		2012Q4 (Two Years Prior)	
	Detected	Other	Detected	Other	Detected	Other
Low	81.41%	82.58%	81.86%	82.58%	81.56%	81.68%
High	18.59%	17.42%	18.14%	17.42%	18.44%	18.32%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

Figure 1 ranks the Detected and Other cases in the descending order of short-term risk. The lower ventiles on the X-axis are associated with higher risk; as one moves left to right on the X-axis, the level of risk goes down. The scale of the Y-axis is the cumulative identification rate. The identification rate is defined as the number of Detected or Undetected cases in each ventile. We compute the identification rates separately for both detected and undetected cases. Based on this graph, it appears that the identification rate of Detected

cases is slightly better than the Other cases, but the difference does not appear to be material. For instance, 14.2 percent of the Detected cases are within the top two ventiles of highest risk. Similarly, 13.4 percent of the Other cases are within the same range. Therefore, the short-term credit score doesn't provide substantial insight in distinguishing noncompliant employers.

FIGURE 1. Identification Rate of Detected and Other Cases in 2014Q4 Based on Short-Term Credit Score



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Undetected	1.9	5.5	9.0	13.4	17.6	24.5	29.5	34.3	40.2	45.2	50.5	56.4	62.8	69.3	75.0	80.7	85.7	90.6	96.3	100.
Detected	2.6	6.7	10.1	14.2	18.6	26.9	32.6	37.8	43.4	49.0	53.9	59.2	64.9	70.7	76.0	80.9	85.2	89.8	96.3	100.

Ventiles of the Population (Ordered from High Risk to Low Risk)

SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

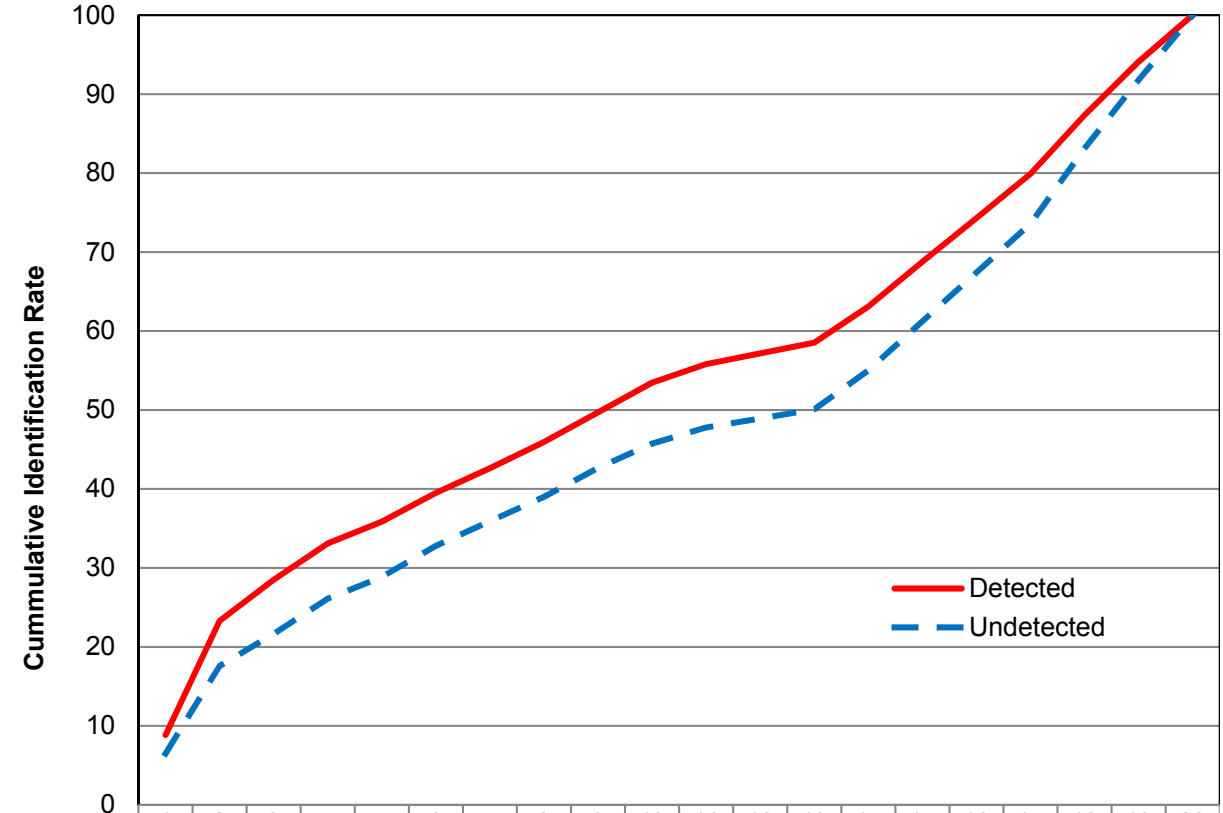
Similarly, Table 3 profiles the Detected and Other cases against the long-term credit score. The identification rate of the Detected cases is higher than that of the Other cases, but there is no clear correspondence between high risk and Detected cases per se.

TABLE 3. Distribution of Low and High Long-Term Credit Score Over Time

Item	2014Q4 (Observation Period)		2013Q4 (One Year Prior)		2012Q4 (Two Years Prior)	
	Detected	Other	Detected	Other	Detected	Other
Low	76.61%	82.58%	75.71%	81.68%	73.76%	79.28%
High	23.39%	17.42%	24.29%	18.32%	26.24%	20.72%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

FIGURE 2. Identification Rate of Detected and Other Cases in 2014Q4 Based on Long-Term Credit Score



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Detected	8.8	23.3	28.5	33.1	35.9	39.5	42.6	45.9	49.7	53.5	55.8	57.2	58.5	63.1	68.8	74.3	79.9	87.3	94.1	100.
Undetected	6.5	17.6	21.7	26.1	28.9	32.8	35.9	39.0	42.7	45.7	47.8	48.9	50.1	55.0	61.3	67.4	73.6	83.1	91.8	100.

Ventiles of the Population (Ordered from High Risk to Low Risk)

SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

Figure 2 is similar to Figure 1, as it represents the identification rates of detected noncompliant employers and other employers with long-term credit risk. The graph suggests that the long-term credit score provides substantial insight in distinguishing the Detected cases from the Other cases because there is a clear separation of the two lines in Figure 2. For instance, 33.1 percent of the Detected cases are within the top two ventiles of highest risk, but just 26.1 percent of the Other cases lie within the same range.

Next, we study whether using both the short- and long-term credit scores results in better identification of different risk categories than studying them separately. Tables 4A and 4B report the Detected and Other cases, respectively, in the same risk category matrix for the observation period 2014Q4.

Tables 4A and 4B show that a higher percentage of Detected cases are in the High Risk and Medium Risk segments compared to the Other cases. Conversely, a higher percentage of the Other cases fall in the Low Risk and Slow Recovery segments compared to the Detected group. Based on these two tables, it appears that the application of both scores simultaneously can provide better identification of potential payroll noncompliance.

TABLE 4A. Distribution of Detected Cases Across Two Credit Scores Concurrently for 2014Q4

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (67.9%)	Medium Risk* (13.5%)
High	Slow Recovery** (8.8%)	High Risk (9.8%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

TABLE 4B. Distribution of Other Cases Across Two Credit Scores Concurrently for 2014Q4

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (72.4%)	Medium Risk* (10.0%)
High	Slow Recovery** (10.1%)	High Risk (7.5%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

For Table 5, because 66.7 percent of the sample used for this paper is comprised of employers with a detected payroll tax underpayment, a percentage higher than 66.7 percent in any of the cells will imply that the application of credit risk scores results in a better identification of detected cases. Table 5 reports this computation for the 2014Q4 period.

TABLE 5. Detected Noncompliance Rates by Risk Category for 2014Q4

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (65.3%)	Medium Risk* (73.0%)
High	Slow Recovery** (63.7%)	High Risk (72.0%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

The same information from Table 5 is transformed into Table 6A by netting out the overall detection rate from each cell. A positive net percentage suggests improvement in identification due to the use of credit risk scores.

TABLE 6A. Net Improvement of Detected Noncompliance Rates by Risk Category for 2014Q4

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (-1.4%)	Medium Risk* (6.3%)
High	Slow Recovery** (-3.0%)	High Risk (5.3%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

The results on Table 6A suggest that conjoint application of the short- and long-term credit risk scores helps in better identification of Medium- and High-Risk segments. The improvement in identification ranges from 5.3 percent to 6.3 percent. The application of both of the credit risk scores is important to clearly identify businesses that belong to different risk categories and each risk category requires specific treatment since businesses in each category face different challenges (Experian, 2016b). Application of long-term credit risk score only segments the sample into low and high risk categories but fails to fully identify the risk segments namely, stable, slow recovery, medium, and high risk. This limitation is circumvented by applying the short-term risk score in conjunction with the long-term risk score. By only applying the long-term credit risk score the low risk category is under identified by -1.6 percent and the High-Risk category is over identified by 5.9 percent. Whereas, applying both the risk scores, the Medium- and the High-Risk segments are not only identified independently but also overidentified by 6.3 and 5.3 percent, respectively. It is important to segregate the Stable segment from Slow Recovery segment since the latter categories require more urgent attention and potential intervention.

Tables 6B and 6C report the same analysis as shown in Table 5, but for the quarters 2013Q4 and 2012Q4, that is, one and two years prior to the observation point, respectively.

TABLE 6B. Net Improvement of Detected Noncompliance Rates by Risk Category for 2013Q4

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (-1.5%)	Medium Risk* (6.3%)
High	Slow Recovery** (-2.8%)	High Risk (5.3%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

TABLE 6C. Net Improvement of Detected Noncompliance Rates by Risk Category for 2012Q4

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (-1.5%)	Medium Risk* (6.3%)
High	Slow Recovery** (-3.3%)	High Risk (4.8%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

The results from Tables 6B and 6C suggest that the conjoint risk scores identify the Medium-Risk and High-Risk cases better than the observed overall rate of 66.7 percent and much earlier than the observation point. Applying only the long-term risk score to the analysis presented in Table 6B, under identifies the Low-Risk category by 1.7 percent and over identifies the High-Risk category by 5.7 percent. Similarly, on Table 6C, application of the long-term risk score over identifies the High-Risk category by only 5 percent but under identifies the Low-Risk segment by 1.6 percent. However, in both the cases, the long-term High-Risk category fails to distinguish between Medium-Risk and High-Risk segments. This demarcation is important since Medium- and High-Risk segments have very different characteristics and they need different types of intervention or treatment. Earlier detection of potentially noncompliant cases may help the IRS recover unpaid payroll taxes, prevent the accumulation of further unpaid payroll taxes, provide guidance to potentially noncompliant employers on the importance of remaining tax compliant, and information on payment options that will help them stay in compliance.

Tables 7, 8, and 9 examine whether the application of both the short- and long-term credit scores help in identifying potential payroll noncompliance in specific segments (legal issues, credit card balance greater than $5,000, and age of business) of the Detected cases.

TABLE 7. Net Improvement of Detected Noncompliance Rates for Cases Having Legal Issues for 2014Q4†

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (-2.8%)	Medium Risk* (-4.1%)
High	Slow Recovery** (8.3%)	High Risk (20.4%)

† The percentages in parentheses represent the net percentage of detected noncompliance cases with legal issues in excess of the overall rate of 24.5 percent.
* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

In Table 7, legal issues are defined as the existence of tax liens at Federal, State and local tax levels, bankruptcies, credit accounts in collections, and Uniform Commercial Code (UCC) filings on a business as of 2014Q3. Based on this analysis, application of both the short- and long-term credit scores can help identify potential payroll noncompliance among employers with legal issues by 8.3 percent and 20.4 percent respectively for the Slow Recovery and High-Risk segments of the sample. In Table 7, applying only the short-term risk score (and ignoring the long-term risk score) would result in under-identification of the short-term low risk segment by 3 percent and over-identification of the short-term High-Risk segment by 14.3 percent. Moreover, this classification results in combining the Stable and Medium-Risk segments into one Low-Risk category and the Slow Recovery and High-Risk segments into one High-Risk category. This combination of dissimilar categories into one category may not result in efficient intervention and treatment since their characteristics are inherently different.

Table 8, examines Detected cases with an average credit card balance of $5,000 across all the credit channels at the reference point 2014Q4. The results indicate that this analysis could identify potential payroll noncompliance among employers in the Slow Recovery group. This is an important segment because businesses in this category are working to improve their credit rating and may be very receptive to outreach and education about the importance of compliance and payment options.

TABLE 8. Net Improvement of Detected Noncompliance Rates for Cases Having Average Balance of $5,000 Across All Credit Lines for 2014Q4[3]

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (1.8%)	Medium Risk* (-11.3%)
High	Slow Recovery** (14.8%)	High Risk (-3.4%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

[3] The percentages in parentheses represent the net percentage of detected noncompliance cases with an average balance of $5000 in excess of the overall rate of 11.3 percent.

Table 9 studies the association of the credit risk scores with the age of the business. We hypothesize that the newest businesses (less than 3 years old) are more vulnerable to lower credit scores and potential tax noncompliance. This table suggests that application of both credit scores helps in identifying potential noncompliance among new businesses that are in Medium-Risk and High-Risk categories. In Table 9, using just the long-term risk score results in the under-identification of the Low-Risk category by 1.7 percent and over-identification of the High-Risk category by 6.2 percent. But using only the long-term risk score fails to distinguish between the Medium-Risk and High-Risk segments, which have very distinct characteristics. Similarly, applying just the short-term risk score results in under-identification of the Low-Risk and over-identification of the High-Risk categories by 0.3 and 1.5, respectively. This result suggests that it is important to use both risk scores since they are complementary to each other, representing different risk segments.

TABLE 9. Net Improvement of Detected Noncompliance Rates for Businesses With Age Less Than 3 Years for 2014Q4[4]

Short-Term Risk	Long-Term Risk	
	Low	High
Low	Stable Segment (-1.7%)	Medium Risk* (7.5%)
High	Slow Recovery** (-1.2%)	High Risk (4.4%)

* Potential risk of financial instability in the long run.
** Slow payment on credit obligations, but chances of surviving through financial hardship is positive.
SOURCE: RAAS Taxpayer Behavior Lab, May 2017.

Conclusion and Future Research

This paper performs a preliminary analysis of short- and long-term credit scores related to detected noncompliant employers vs. other employers. It provides exploratory evidence that concurrent application of both the short- and long-term credit scores may help in identifying potential payroll tax noncompliance. This may help the IRS in understanding the behavior of these businesses, enable the IRS to take proactive steps to recover unpaid payroll taxes, and educate employers about payment plans that can help them meet their payroll tax obligations.

An extension of this paper will include examining the association between changes in credit score and noncompliance by matching credit bureau data with IRS administrative data. Furthermore, we plan to examine whether a change from the Low-Risk category to the High-Risk category is associated with future noncompliance. A comprehensive understanding of the causality between the two credit scores may also help in understanding how credit risk scores may help in identifying noncompliance with suitable lags after controlling for other factors in an econometric model. Another extension will be to use the Markov Transition Matrix from the credit risk modeling literature. A Transition Matrix structure (Jones (2005); Dobrow (2016)) can be employed to study the relationship between transition of credit risk categories and potential future noncompliance. This structure will help in detecting whether the initial credit risk state or movement from different risk categories in the past is associated with future employment tax noncompliance. Appropriate techniques need to be employed to identify the optimum look back period to effectively predict the timing of potential future noncompliance.

[4] The percentages in parentheses represent the net percentage of detected noncompliance cases (with age of the business being less than 10 years) in excess of the overall rate of 11.7 percent.

References

Dobrow, R. (2016). *Introduction to Stochastic Processes with R*. Wiley.

Experian. (2016a). *Experian Intelliscore Plus V2*. Retrieved from Experian.com: https://www.experian.com/assets/business-information/brochures/intelliscore-plus-v2-product-sheet.pdf.

Experian. (2016b). *Experian Stability Risk Score*. Retrieved from Experian.com: https://www.experian.com/assets/business-information/brochures/financial-stability-risk-score-ps.pdf.

Jones, M. (2005). *Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk*. International Monetary Fund, Monetary and Financial Systems Department. IMF.

SB/SE Research New Carrollton/Richmond. (2011). *Benefit and Cost Evaluation of the Federal Tax Deposit Alert (FTD) Program* (Vol. NCH0149).

Hopkins, C. and K. Su (2017). Supplementing IRS Data with External Credit Report Data in Employment Tax Predictive Models. *IRS Research Bulletin* (elsewhere in this volume).

Estimating the Effects of Tax Reform on Compliance Burdens

Daniel Berger and Eric Toder (Urban Institute-Tax Policy Center), and Victoria Bryant, John Guyton, and Patrick Langetieg (IRS Research, Applied Analytics and Statistics)[1]

Introduction

The primary purpose of the tax code is to collect tax revenue; however, the complexity of the tax code imposes costs on taxpayers beyond what they are paying in taxes. These costs include time spent on record keeping, planning and gathering materials, and out-of-pocket outlays for tax software, paid-preparer services and other associated fees.[2] These real economic costs to tax filers due to tax code complexity should be counted when analyzing the effects of tax reform proposals.

The Tax Policy Center (TPC), in conjunction with the Internal Revenue Service Office of Research, Applied Analytics and Statistics (RAAS), has recently developed distributional estimates of individual compliance costs for tax filers. This paper describes the process used by TPC to incorporate the model used by RAAS into TPC's Individual Income Tax Microsimulation Model, and presents estimates produced by TPC's model.[3]

We begin with a brief discussion of microsimulation modeling, followed by a description of the burden model used by RAAS. The paper then explains how the IRS altered their model to make it workable with a public data-based microsimulation model. Finally, we present compliance cost estimates produced by TPC under current law and three tax reform options. All the estimates are based on modifications to the tax law that existed prior to enactment of the Tax Cut and Jobs Act of 2017 (TCJA).

Tax Simulation Outside the Government

The IRS Statistics of Income (SOI) division each year creates a stratified random sample of individual income tax returns to aid Executive Branch and Congressional agencies in the development and analysis of tax policies. A full description of the SOI sample can be found in Testa and Haines (2012). By statute, the only organizations allowed access to such confidential data outside of the IRS are the Office of Tax Analysis (OTA) at the Treasury Department and the Joint Committee on Taxation (JCT).[4] Some researchers who are bound by confidentiality agreements also have limited access to these data in secure facilities for specified research purposes intended to assist in the administration of the tax law.

The IRS also produces a public use version of the SOI sample, known as the Public Use File (PUF), by drawing a subsample from the full sample. To ensure taxpayer confidentiality, the PUF incorporates several nondisclosure procedures that remove direct and indirect disclosure of individual taxpayers. These procedures include omitting taxpayer identifying information, removing sensitive returns, further subsampling high income returns, and limiting the tax return fields released, among others. A more comprehensive description of these procedures can be found in Bryant (2016) and Bryant, *et al.* (2016). The PUF is designed to replicate as closely as possible the statistical results from the larger sample, with the qualification that some aggregate data cannot be generated due to the missing fields.

---

[1] Prepared for the 2017 IRS-TPC Research Conference. Views expressed are those of the authors and should not be attributed to funders or affiliated institutions. The authors would like to thank Barry Johnson, Allen Lerman, Emily Lin, Joe Rosenberg, Jeff Rohaly and Alan Plumley for their comments on earlier drafts of this work.

[2] See Marcuss, *et al.* (2013).

[3] All model estimates are based on law before the enactment of the Tax Cuts and Jobs Act of 2017.

[4] The JCT has also allowed the Congressional Budget Office (CBO) to use the database to aid in developing revenue forecasts and at times the Government Accountability Office (GAO) has received access to enable its audits of tax policy and administration.

Individuals or organizations can purchase the PUF, but must then invest substantial programming resources to develop tax calculators and other programs to generate statistical information on the taxpaying population. Over the years, most users of the PUF have been universities, other research institutions, accounting firms, or government agencies. The CBO began using the PUF in the 1980s and still uses it for tax research studies not directly related to developing economic and revenue projections. Another major user over the years has been the National Bureau of Economic Research, which has produced many studies using its TAXSIM model.

The PUF is a rich data source, but it does not contain key information of interest to researchers that is not reported on tax returns. Therefore, researchers have supplemented tax return data in the PUF with household demographic data from other micro data files, including the Current Population Survey (CPS) produced by the U.S. Census Bureau and the Survey of Consumer Finances (SCF) produced by the Federal Reserve Board.

The Tax Policy Center (TPC), founded in 2002, developed its individual tax model based on the PUF. The current version of the model starts with the 2006 PUF, which is a sample of slightly under 150,000 tax returns. The 2006 data are adjusted by the growth of income and deduction items to match reported 2011 data by filing status and income group published by SOI. This adjustment creates what TPC calls the 2011 Look Alike Public Use File (LAPUF), which combines individual taxpayer variation from the 2006 PUF with aggregate SOI published data from 2011.

TPC uses cross-tabulations by age, filing status, and income provided by SOI to impute the ages of taxpayers and dependents to the LAPUF. TPC then performs a constrained statistical match between the LAPUF and the 2012 CPS. The CPS match adds additional demographic information to the file and allows TPC to augment the sample of tax return filers with nonfilers from the CPS to create a more complete sample of the U.S. population, grouped into "tax units." TPC then augments the matched data set with imputations of consumption, health insurance status and benefits, pension coverage and assets, and other demographic variables from a variety of data sources. The data file is then aged to years 2012 through 2027, based on CBO and JCT baseline revenue projections, JCT estimates of the distribution of tax units by income, IRS estimates of the future growth in the number of tax returns, and Census projections of the size and age-composition of the population.

A more complete description of the database that TPC has developed and the methods it uses to estimate federal receipts and the distribution of the tax burden from different federal tax sources can be found at: http://www.taxpolicycenter.org/resources/brief-description-tax-model.

Background on Burden Model

Guyton, *et al.* (2003) and Marcuss, *et al.* (2013) provide an overview of the principles and methods underlying the IRS Compliance Burden Models. The rationale for developing compliance burden microsimulation modeling capabilities to support tax administration and policy analysis dates back to a task force established in 1998 by IRS Commissioner Rossotti with support from the Treasury Office of Tax Analysis, the Office of Management and Budget, and the Government Accountability Office, as well as academic advisors.

The Individual Taxpayer Burden Model (ITBM) was first put into official use to support Paperwork Reduction Act (PRA) reporting starting with Tax Year 2005. The modeling approach was refined into an early version of its current specification as part of the effort to extend the framework to cover corporate and partnership taxpayers in 2009, influenced by earlier work on the compliance costs of large corporations by Slemrod and Blumenthal (1996) and by Slemrod and Venkatesh (2002). Official PRA reporting using the Business Taxpayer Burden Model (BTBM) started with Tax Year 2014, based on the specification described in Marcuss, *et al.* (2013). The current specification of the ITBM also dates from Marcuss, *et al.* (2013), but has been updated since then using more recent taxpayer burden survey data. Additional information on the development and use of the taxpayer burden model can be found in Contos, *et al.* (2009a), Contos, *et al.* (2009b), Contos, *et al.* (2010), Contos, *et al.* (2012), and the 2013 Economic Report of the President.[5]

[5] Council of Economic Advisers (2013).

Adapting the Individual Taxpayer Burden Model for the Public Use File

Adapting the Individual Taxpayer Burden (ITB) model for the PUF requires adjustments to the ITB model inputs and updated regression coefficients. The adjustments are needed to ensure the model inputs are consistent with the level of data granularity available on the PUF. Once the adjustments are made, consistent model coefficients can be estimated and then applied to the PUF to estimate taxpayer burden. Appendix 1 provides details about how the complexity category model inputs were redefined to be consistent with the level of granularity available on the PUF. Technical documentation about the burden estimation process using the PUF is provided in an electronic technical Appendix.[6] More detail about the ITB model specification and estimation procedures can be found in Contos, *et al.* (2009a), Contos, *et al.* (2009b), Contos, *et al.* (2010), and Marcuss, *et al.* (2013).

Tables 1–5 in the following section present summary statistics and simulation results that compare the similarities and differences before and after PUF adjustments. The tables include a comparison of the data used in the regression specification, regression coefficients, and distributions of simulated burdens. The results presented here use the data from the Tax Year 2007 individual taxpayer burden survey study. Results for other survey years can be found in Appendix Tables 2–5.

The ITB Model Before and After the PUF Adjustments

Differences in the Data

The PUF adjustments affect only the continuous fields used in the ITB model specification. Table 1 provides descriptive statistics for the continuous fields before and after the PUF adjustments.

TABLE 1. Model Input Distributional Statistics, Percentile Breakdown of Model Inputs, 2007

Field Description	Mean	P5	P10	P25	P50	P75	P90	P95
Before PUF Adjustments								
Log Modified Positive Income	10.42	8.39	8.93	9.72	10.50	11.21	11.77	12.16
Log of Low Complexity Activity	45.15	8.82	15.76	27.04	40.69	62.16	79.57	88.26
Log of Medium Complexity Activity	25.78	0.00	0.00	0.00	10.37	38.21	72.46	97.20
Log of High Complexity Activity	1.45	0.00	0.00	0.00	0.00	0.00	7.22	10.54
Log of Line Count	4.70	2.56	2.89	4.36	4.88	5.29	5.64	5.81
After PUF Adjustments								
Log Modified Positive Income	10.37	8.37	8.91	9.70	10.48	11.18	11.71	12.08
Log of Low Complexity Activity	39.70	8.13	14.93	26.07	36.62	52.75	66.82	74.96
Log of Medium Complexity Activity	18.24	0.00	0.00	0.00	9.60	28.64	49.25	63.18
Log of High Complexity Activity	0.69	0.00	0.00	0.00	0.00	0.00	0.00	7.49
Log of Line Count	4.68	2.56	2.89	4.36	4.84	5.26	5.58	5.74
Percent Change from PUF Adjustments								
Log Modified Positive Income	-0.49%	-0.32%	-0.25%	-0.21%	-0.14%	-0.23%	-0.54%	-0.65%
Log of Low Complexity Activity	-12.08%	-7.85%	-5.29%	-3.57%	-10.00%	-15.14%	-16.03%	-15.07%
Log of Medium Complexity Activity	-29.26%	0.00%	0.00%	0.00%	-7.50%	-25.03%	-32.03%	-35.00%
Log of High Complexity Activity	-52.70%	0.00%	0.00%	0.00%	0.00%	0.00%	-100.00%	-28.95%
Log of Line Count	-0.47%	0.00%	0.00%	0.00%	-0.64%	-0.68%	-1.04%	-1.17%

[6] The technical appendix can be found at https://www.irs.gov/statistics/irs-taxpayer-compliance-research.

The PUF data limitations do not appear to have a material impact on the modified positive income measure. This outcome seems reasonable because modified positive income is highly correlated with total income and typically deviates from total income only when taxpayers have gains and losses on multiple Schedule Cs or Fs.

The PUF data limitations do appear to have a material impact on the complexity category measures. This observation is also in line with expectations. The complexity measures are constructed as the sum of the logs of line item amounts and the PUF is missing several line items on supporting forms and schedules, so the observed decrease is expected and unavoidable. Some of these missing line items can be partially counted using residual portions of summation lines, such as total income or total adjustments, but several other line items are simply missing.

Differences in the Regression Parameters

Table 2 provides regression parameter estimates using the Tax Year 2007 ITB survey data before and after PUF adjustments. The dependent variable is the log of burden where burden is defined as out of pocket expenses plus monetized time.

TABLE 2. Regression Coefficients*

Model Details	Adjusted Regression Parameters	
	Before PUF Adjustment	After PUF Adjustment
Intercept	0.4788	0.4944
Log Modified Positive Income | Used a Paid Preparer	0.1806	0.1646
Log Modified Positive Income | Used a Tax Software	0.2350	0.2336
Log Modified Positive Income | Prepared by Hand	0.2516	0.2482
Log of Low Complexity Activity	0.0085	0.0089
Log of Medium Complexity Activity	0.0081	0.0148
Log of High Complexity Activity	0.0228	0.0233
Log of Line Count | Used a Paid Preparer	0.0396	0.0812
Log of Line Count | Used Tax Software	0.1678	0.1706
Log of Line Count | Prepared by Hand	0.2846	0.2845
Used a Paid Preparer	2.5758	2.5344
Used Tax Software	0.9530	0.9399
Adjusted R-Squared	0.5644	0.5666

* Coefficients based on the ITB Survey, 2007.

All the regression coefficients have the same sign before and after the PUF adjustment and most of them change little in magnitude. The coefficients for the complexity categories are all larger after the PUF adjustment. This increase seems reasonable given the drop in the complexity category measures observed in Table 1.

The most significant difference between the two models is the change in the "Log of Line Count | Used a Paid Preparer" coefficient. Part of the change is likely offset by the decrease in the "Log Modified Positive Income | Used a Paid Preparer" coefficient. Conceptually, this coefficient is still consistent. For example,

taxpayers choosing to use a paid preparer have the highest fixed costs and the lowest variable costs while taxpayers choosing to prepare without the help of a paid preparer or software have the lowest fixed costs and the highest variable costs. Overall, the coefficients are slightly different so the resulting distribution of predicted burden is expected to be slightly different as well. The next section will evaluate the distributional differences between the estimates from the two models.

Differences in Burden Simulation Results

Table 3 provides a set of distributional results using the 2007 ITB survey sample data. The table provides the distribution of burden as reported on the survey instrument, the distribution of burden estimated using the model before PUF adjustments, and the distribution of burden estimated using the model after PUF adjustments.

TABLE 3. Distributional Statistics for Reported and Estimated Total Tax Burden, 2007

Percentile of the Distribution of Burden	Burden per Tax Unit as Reported ($)	Estimated Burden Before PUF Adjustment ($)	Reported vs Before Adjustment Percent Change	Estimated Burden After PUF Adjustment ($)	Reported vs After Adjustment Percent Change
99%	$3,089	$3,252	5.28%	$3,086	-0.10%
95%	$1,332	$1,293	-2.93%	$1,299	-2.48%
90%	$845	$828	-2.01%	$832	-1.54%
75%	$394	$404	2.54%	$405	2.79%
50%	$181	$183	1.10%	$181	0.00%
25%	$77	$79	2.60%	$79	2.60%
10%	$32	$36	12.50%	$35	9.38%
5%	$20	$22	10.00%	$22	10.00%
1%	$9	$9	0.00%	$8	-11.11%
Mean	$371	$387	4.31%	$380	2.43%

Distributions for estimated burden before PUF adjustments and after PUF adjustments differ only slightly from the underlying survey data. They also appear to track each other closely, differing from the survey instrument distribution in roughly the same places. The percentage change columns show that the after PUF adjustment model is as good or better than the before PUF adjustment model at recreating the distribution of burden from the survey instrument. This suggests that the adjusted ITB model inputs and the regression coefficients using the PUF provide estimates that are roughly consistent or better than the estimates using the unadjusted ITB model.

Table 4 provides the same distributional information as in Table 3, but broken out by preparation method.

TABLE 4. Distributional Statistics for Reported and Estimated Total Taxpayer Burden by Preparation Method, 2007

Percentile of the Distribution of Burden	Burden per Tax Unit as Reported ($)	Estimated Burden Before PUF Adjustment ($)	Reported vs Before Adjustment Percent Change	Estimated Burden After PUF Adjustment ($)	Reported vs After Adjustment Percent Change
Used a Paid Preparer					
99%	$3,816	$4,209	10.30%	$3,957	3.69%
95%	$1,653	$1,664	0.67%	$1,658	0.30%
90%	$1,042	$1,067	2.40%	$1,076	3.26%
75%	$507	$534	5.33%	$537	5.92%
50%	$262	$260	-0.76%	$258	-1.53%
25%	$129	$129	0.00%	$127	-1.55%
10%	$69	$69	0.00%	$68	-1.45%
5%	$42	$48	14.29%	$47	11.90%
1%	$21	$25	19.05%	$23	9.52%
Mean	$488	$517	5.94%	$508	4.10%
Prepared by Hand					
99%	$1,424	$937	-34.20%	$929	-34.76%
95%	$471	$409	-13.16%	$406	-13.80%
90%	$284	$265	-6.69%	$260	-8.45%
75%	$123	$129	4.88%	$126	2.44%
50%	$54	$58	7.41%	$57	5.56%
25%	$23	$27	17.39%	$26	13.04%
10%	$12	$13	8.33%	$12	0.00%
5%	$9	$8	-11.11%	$8	-11.11%
1%	$4	$4	0.00%	$4	0.00%
Mean	$127	$117	-7.87%	$115	-9.45%
Used Tax Software					
99%	$1,873	$1,816	-3.04%	$1,765	-5.77%
95%	$961	$847	-11.86%	$849	-11.65%
90%	$598	$567	-5.18%	$572	-4.35%
75%	$279	$292	4.66%	$289	3.58%
50%	$131	$135	3.05%	$133	1.53%
25%	$56	$62	10.71%	$61	8.93%
10%	$29	$31	6.90%	$30	3.45%
5%	$19	$20	5.26%	$19	0.00%
1%	$9	$9	0.00%	$8	-11.11%
Mean	$254	$253	-0.39%	$250	-1.57%

These simulation results show that the full sample and PUF distributional estimates track each other closely and differ from the survey instrument distribution in roughly the same places for each preparation method. This is further evidence that the adjusted ITB model inputs and regression coefficients using the PUF provide estimates that appear to be consistent with the unadjusted ITB model.

TPC and IRS Models' Composition of Burden

Table 5 shows differences in the distribution of compliance costs by income and tax categories between the IRS full model, and the TPC PUF-based model. The TPC model was calibrated to match aggregate totals of the IRS full model, but does not necessarily match totals for each source of income and income offset. The differences for separate sources of income and adjustments can at least partially be explained by the different data sets used for the analysis.

TABLE 5. Composition of Burden Cost: Percent of Total Burden Cost, 2017[a]

Income / Tax Sources	TPC (%)	IRS (%)
Other Taxes and Items Not Related to Income Tax	6.33	4.00
Alternative Minimum Tax (AMT)	4.76	2.00
Credits	12.65	14.00
Deductions	29.38	25.00
Self-Employment Income	14.37	19.00
All Other Nonwage Income	11.25	18.00
Wages	21.26	18.00
Total	**100.00**	**100.00**

NOTES: (a) Calendar year estimates. Estimates are derived by removing income and tax sources from compliance cost in the order presented in figure.
SOURCE: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1); Economic Report of the President, March 2013, Figure 3–10; https://obamawhitehouse.archives.gov/sites/default/files/docs/erp2013/full_2013_economic_report_of_the_president.pdf.

Table 5 shows the composition of burden in the IRS full model and the TPC model. The order of the table represents the stacking order, or the order items were removed from the burden model to estimate the percent composition of burden. Because certain items available to the IRS are missing from the PUF, we see variation between the composition of burden in the IRS and TPC models. This can be seen in the All Other Non-Wage Income category. In this category, both the IRS and TPC removed Schedule C income (along with other items), but as the PUF has fewer Schedule C items, the effect of removing All Other Non-Wages for TPC (11%) is smaller than for IRS (18%). Though most of the differences are relatively small, it should be noted that they could affect the allocation of burden, thus causing distortions for estimates that repeal or modify sections of the tax code.

Distributional Analysis of Compliance Burden

TPC has estimated the total compliance burden of individual taxes and its distribution by expanded cash income (ECI) group in the current tax system, along with the changes in compliance cost by income group for three revenue neutral tax changes.

TABLE 6. Distribution of Baseline Compliance Cost by Expanded Cash Income Percentile, 2017[a]

Expanded Cash Income Percentile[b,c]	Tax Units		Percent Change in After Tax Income[d]	Share of Total Federal Tax Change	Average Federal Tax Change ($)	Average Federal Tax Rate[e]	
	Number (Thousands)	Percent of Total				Change (Percentage Points)	Under the Proposal
Quintiles							
Lowest	48,560	27.8	-0.9	5.4	116	0.8	5.0
Second	38,510	22.0	-0.8	9.2	249	0.7	9.5
Middle	34,010	19.4	-0.8	14.2	436	0.7	14.6
Fourth	28,660	16.4	-0.8	19.8	720	0.6	18.2
Top	24,130	13.8	-0.9	51.4	2,218	0.7	26.4
All	174,990	100.0	-0.8	100.0	596	0.7	20.7
Other							
80–90	12,380	7.1	-0.8	13.6	1,147	0.6	21.1
90–95	5,990	3.4	-0.9	9.8	1,703	0.7	22.9
95–99	4,630	2.6	-1.0	14.6	3,277	0.8	26.5
Top 1%	1,140	0.7	-0.9	13.4	12,270	0.6	33.5
Top 0.1%	120	0.1	-0.8	5.9	53,319	0.5	34.6

NOTES:
(a) Calendar year. Baseline is current law, with no assumed compliance costs. Estimates add compliance costs into the economy. For a description of TPC's current law baseline, see http://www.taxpolicycenter.org/resources/tpc-baseline-definitions.
(b) Includes both filing and nonfiling units but excludes those that are dependents of other tax units. Tax units with negative adjusted gross income are excluded from their respective income class but are included in the totals. For a description of expanded cash income, see http://www.taxpolicycenter.org/TaxModel/income.cfm.
(c) The income percentile classes used in this table are based on the income distribution for the entire population and contain an equal number of people, not tax units. The breaks (in 2017 dollars) are: 20% $28,100; 40% $54,700; 60% $93,200; 80% $154,900; 90% $225,400; 95% $304,600; 99% $912,100; 99.9% $5,088,900.
(d) Includes tax units with a change in Federal tax burden of $10 or more in absolute value.
(e) Average Federal tax (includes individual and corporate income tax, payroll taxes for Social Security and Medicare, the estate tax, and excise taxes) as a percentage of average expanded cash income.
SOURCE: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1).

Table 6 shows TPC's baseline estimates for the compliance cost of filing individual taxes. We estimate that the cost of individual compliance for Calendar Year 2017 is $596 dollars per tax unit for 174.9 million units, making the total cost equal to just over $104 billion.

These estimates are compared to a baseline scenario with an assumption of no individual compliance cost, so for the purposes of Table 6, compliance costs are effectively being presented as if they were a legislated tax increase.

Tax units in all income groups incur compliance costs associated with filing taxes. Tax units in higher ECI quintiles incur higher absolute compliance costs per tax unit, but compliance costs rise less than proportionately with income through the four bottom, quintiles before increasing through the bottom 95 percent of the top quintile (the 80–99th percentiles of the distribution). The top quintile, with slightly over half (52.1 percent) of pretax income, incurs slightly over half the total compliance costs (51.4 percent), while the bottom income quintile, with 4.4 percent of pretax income, incurs 5.4 percent of the total compliance burden. Average compliance costs as a share of pretax income decrease from 0.8 percent for the bottom quintile to 0.7 percent for the second and third quintile, and 0.6 for the fourth quintile, increasing back to 0.7 for the top quintile. Within the top quintile, costs as a share of income increase slightly to 0.8 percent for tax units in the 95th to 99th percentiles before declining to 0.6 percent in the top 1 percent and 0.5 percent for the top 0.1 percent.

Hypothetical Changes in Tax Law

Tables 7–9 below include current law compliance costs in the baseline measure of taxes paid, and display the effects of changes in tax law on the sum of tax liability and compliance costs for three proposed tax changes.

1. Repeal all itemized deductions and increase the standard deduction

The first option (Table 7) would repeal all itemized deductions and use the revenues raised to finance an increase in the standard deduction. To achieve revenue neutrality, we multiplied the standard deduction for each filing status, the additional deduction for the elderly and blind and the dependent standard deduction by 2.13.

TABLE 7. Changes in Taxes Paid and Compliance Costs of Option 1 (as a share of pretax income), 2017[a]

Expanded Cash Income Percentile[b,c]	Tax	Compliance	Tax & Compliance
Quintiles			
Lowest	-0.5	-0.2	-0.7
Second	-1.4	-0.2	-1.6
Middle	-1.5	-0.2	-1.7
Fourth	-1.1	-0.2	-1.3
Top	1.2	-0.3	0.9
All	0.0	-0.2	-0.2
Other			
80–90	-0.3	-0.2	-0.6
90–95	0.7	-0.3	0.5
95–99	1.6	-0.3	1.3
Top 1%	2.5	-0.3	2.2
Top 0.1%	2.9	-0.3	2.6

NOTES:
(a) Calendar year. Baseline is current law with associated compliance costs of Option 1. Proposal would repeal all itemized deductions and proportionally increase the standard deduction by a factor of 2.13 to achieve revenue neutrality. For a description of TPC's current law baseline, see http://www.taxpolicycenter.org/resources/tpc-baseline-definitions.
(b) Includes both filing and nonfiling units but excludes those that are dependents of other tax units. Tax units with negative adjusted gross income are excluded from their respective income class but are included in the totals. For a description of expanded cash income, see http://www.taxpolicycenter.org/TaxModel/income.cfm.
(c) The income percentile classes used in this table are based on the income distribution for the entire population and contain an equal number of people, not tax units. The breaks (in 2017 dollars) are: 20% $28,100; 40% $54,700; 60% $93,200; 80% $154,900; 90% $225,400; 95% $304,600; 99% $912,100; 99.9% $5,088,900.
SOURCE: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1).

Higher income tax payers are much more likely to choose to itemize deductions than others, and with the proposed repeal, the highest income quintile is the only group to see an average increase in tax burden[7] (Table 7). Taxes as a share of income would increase by 1.2 percent for the highest income quintile and 2.9 percent for the top 0.1 percent. Tax units in the bottom four quintiles would benefit more from the increase in the standard deduction than from the loss of itemized deductions. Taxes would decline for all four quintiles, with the largest benefits going to the middle quintile (1.5 percent of income) and the second quintile (1.4 percent of income).

Compliance costs would decline by 0.2 percent of pretax income for taxpayers in the bottom four income quintiles and by about 0.3 percent of income for the top quintile. The proposal would lower compliance costs for all tax units that would otherwise itemize deductions. Additionally, certain tax units in the lowest

[7] See Lu (2017).

and second quintile would benefit as they may no longer be required to file taxes. Currently individuals are required to file taxes if their gross income is greater than the standard deduction plus their allotted personal exemptions.[8] With the increase in the standard deduction, many tax units who would file under current law would no longer need to file, and under our estimation parameters would no longer have any compliance cost.

Summing up the changes in tax burdens and compliance burdens, the proposal would reduce total burdens in the bottom four quintiles and increase the combined tax and compliance burden for the top quintile. The reduction in compliance costs benefits all groups, but offsets only a portion of the tax increase for the top quintile, leaving the pattern of the distributional change roughly similar to the distributional pattern of the change in tax payments alone.

2. Eliminate most itemized deductions

The second option would eliminate most itemized deductions, but would retain the deductions for mortgage interest and charitable giving (Table 8). To achieve revenue neutrality the standard deduction for all filing units, the additional deduction for the elderly and blind and the dependent standard deduction were all multiplied by 1.90.

TABLE 8. Changes in Taxes Paid and Compliance Costs of Option 2 (as a share of pretax income), 2017[a]

Expanded Cash Income Percentile[b,c]	Tax	Compliance	Tax & Compliance
Quintiles			
Lowest	-0.5	-0.2	-0.7
Second	-1.1	-0.2	-1.3
Middle	-1.2	-0.2	-1.4
Fourth	-0.8	-0.2	-1.0
Top	0.9	-0.2	0.7
All	0.0	-0.2	-0.2
Other			
80–90	-0.1	-0.2	-0.3
90–95	0.7	-0.2	0.5
95–99	1.2	-0.2	1.0
Top 1 Percent	1.7	-0.2	1.5
Top 0.1 Percent	1.7	-0.2	1.5

NOTES:
(a) Calendar year. Baseline is current law with associated compliance costs. Proposal would repeal all itemized deductions except the mortgage interest and charitable deductions and proportionally increase the standard deduction by a factor of 1.90 to achieve revenue neutrality. For a description of TPC's current law baseline, see http://www.taxpolicycenter.org/taxtopics/Baseline-Definitions.cfm.
(b) Includes both filing and nonfiling units but excludes those that are dependents of other tax units. Tax units with negative adjusted gross income are excluded from their respective income class but are included in the totals. For a description of expanded cash income, see http://www.taxpolicycenter.org/TaxModel/income.cfm.
(c) The income percentile classes used in this table are based on the income distribution for the entire population and contain an equal number of people, not tax units. The breaks (in 2017 dollars) are: 20% $28,100; 40% $54,700; 60% $93,200; 80% $154,900; 90% $225,400; 95% $304,600; 99% $912,100; 99.9% $5,088,900.
SOURCE: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1).

The distributional effects are similar to those for eliminating all itemized deductions, with average tax burden increasing only for those in the highest income quintile. While the average tax as a share of pretax income increases for the top quintile under this option, the increase in tax burdens drops from 1.2 to 0.9 percent of income when allowing them to continue to deduct mortgage interest and charitable contributions. The middle

[8] Certain filers are required to file even if they do not meet the gross income requirements. For a comprehensive list of IRS filing requirements see: https://www.irs.gov/publications/p17/ch01.html#en_US_2016_publink1000170388.

[9] http://www.taxpolicycenter.org/briefing-book/who-pays-amt.

(1.2 percent of income) and second (1.1 percent of income) quintiles still receive the largest decrease in taxes as a share of pretax income. The lowest quintile remains unchanged from the previous proposal with a tax cut of 0.5 percent of pretax income.

As in the proposal to eliminate all itemized deductions, compliance costs decline by about 0.2 percent of pretax income for all quintiles. Tax units above the 90th percentile see their compliance costs decrease slightly less as a share of income when the two deductions are maintained than when all itemized deductions are eliminated, falling by 0.2 instead of 0.3 percent of income (Tables 7 and 8).

3. Repeal the AMT and reduce the State & local tax deduction

The final proposal repeals the Alternative Minimum Tax (AMT), and to maintain revenue neutrality, scales down the State and local tax deduction. This proposal allows tax units to deduct only 80 percent of the current law value of the State and local deduction.

The AMT typically affects tax units that are high income, but less so for units at the very top of the income distribution.[9] The State and local deduction benefits tax units that itemize deductions, which again, are generally higher income earners. This can be seen in Table 9, with tax units in the 95–99th percentile disproportionately benefiting from the repeal of the AMT, while the loss of the State and local income deduction is spread more evenly across the middle, fourth and top quintiles.

The plan would increase average taxes by 0.1 percent of pretax income for the middle and fourth income quintiles, while decreasing taxes in the top quintile by 0.1 percent of income (Table 9). Tax units in the 80-90th percentiles and 90–95th percentile would see a slight increase in taxes (0.2 and 0.1 percent of pretax income). Those in the 95–99th percentiles would receive the largest share of the benefit with a tax cut of 0.5 percent of income, while the top 1 percent (0.1 percent) and the top 0.1 percent (0.2 percent) would receive smaller cuts as a share of pretax income. The proposal would have a negligible effect on the bottom two quintiles.

TABLE 9. Changes in Taxes Paid and Compliance Costs of Option 3 (as a share of pretax income), 2017[a]

Expanded Cash Income Percentile[b,c]	Tax	Compliance	Tax & Compliance
Quintiles			
Lowest	0.0	0.0	0.0
Second	0.0	0.0	0.0
Middle	0.1	0.0	0.1
Fourth	0.1	0.0	0.1
Top	-0.1	-0.1	-0.2
All	0.0	-0.1	-0.1
Addendum			
80–90	0.2	0.0	0.2
90–95	0.1	0.0	0.1
95–99	-0.5	-0.1	-0.7
Top 1%	-0.1	-0.1	-0.2
Top 0.1%	-0.2	-0.1	-0.2

NOTES:
Calendar year. Baseline is current law with associated compliance costs. Proposal would repeal the Alternative Minimum Tax (AMT) and scale down the State and local deduction by 20 percent to achieve revenue neutrality. For a description of TPC's current law baseline, see http://www.taxpolicycenter.org/taxtopics/Baseline-Definitions.cfm.

(b) Includes both filing and nonfiling units but excludes those that are dependents of other tax units. Tax units with negative adjusted gross income are excluded from their respective income class but are included in the totals. For a description of expanded cash income, see http://www.taxpolicycenter.org/TaxModel/income.cfm.

(c) The income percentile classes used in this table are based on the income distribution for the entire population and contain an equal number of people, not tax units. The breaks (in 2017 dollars) are: 20% $28,100; 40% $54,700; 60% $93,200; 80% $154,900; 90% $225,400; 95% $304,600; 99% $912,100; 99.9% $5,088,900.

SOURCE: Urban-Brookings Tax Policy Center Microsimulation Model (version 0217-1).

The proposal would have a modest effect on compliance costs at the top of the income distribution, as those returns would no longer have to file the AMT, but would generate negligible compliance savings for the bottom four income quintiles and, within the top quintile, for tax units in the 80-95th percentiles. Among income recipients in the top quintile, only those in the 95-99th percentiles (0.1 percent of income) and top 1 percent (0.1 percent of income) would see their compliance costs decline. When accounting for changes in tax and compliance cost, tax units in the 95–99th percentiles would get the largest share of the benefit with total taxes and compliance costs decreasing by 0.7 percent of pretax income.

Conclusion

Our analysis of compliance costs shows that the complexity of the tax code cost individuals over $104 billion dollars in Tax Year 2017, or an average of $596 per taxpayer. While compliance costs increase for higher income taxpayers, compliance burden is highest as a share of pretax income for those in bottom income quintile.

The $104 billion figure is comprised of monetized time, out-of-pocket costs and other monetized costs of filing individual taxes. While any tax system will inevitably have costs associated with compliance and other sources of efficiency loss, a goal of tax policy should be to limit these inefficiencies, while collecting the necessary amount of revenue with an equitable distribution of the tax burden.[10]

As seen in the reform options presented in this paper, repealing sections of the tax code can lower compliance burdens, which are one component of the resource cost of taxation. These reductions can work to mitigate a portion of tax increases for those who would otherwise face an increased burden under the plan, and increase benefits for those who would already be benefiting from the proposed changes in tax law. This analysis highlights the fact that when considering tax reform proposals, policy makers should consider the real economic costs associated with adding complication into the tax code, and the added resource benefits of simplifying the existing system. While reducing compliance burdens increases the overall efficiency of the system, it may often, as shown in these examples, have little effect on which groups are net winners and losers from a particular tax policy change.

[10] See GAO (2005).

References

Bryant, V. (2016), *General Description Booklet of the 2011 Public Use File.*

Bryant, V., J. Czajka, G. Ivsin, and J. Nunns (2014), "Design Changes to the SOI Public Use File (PUF)," *107th Annual Conference Proceedings,* 2014, National Tax Association.

Contos, George, Ardeshir Eftekharzadeh, Brian Erard, John Guyton, and Scott Stilmar (2009a), "Econometric Simulation of the Income Tax Compliance Process for Small Business." In *Proceedings of the 2009 Winter Simulation Conference,* 2762–2765. Winter Simulation Conference, http://informs-sim.org/wsc09papers/prog09soc.html.

Contos, George, John Guyton, Patrick Langetieg, and Susan Nelson (2009b), "Taxpayer Compliance Costs for Small Businesses: Evidence From Corporations, Partnerships, and Sole Proprietorships." *Proceedings of the One Hundred Second Annual Conference on Taxation,* 50–59. National Tax Association, Washington, DC.

Contos, George, John Guyton, Patrick Langetieg, and Melissa Vigil (2010), "Individual Taxpayer Compliance Burden: The Role of Assisted Methods in Taxpayer Response to Increasing Complexity." In *Proceedings of the 2010 IRS Research Conference,* 191–220. Internal Revenue Service, Washington, DC.

Contos, George, John Guyton, Patrick Langetieg, Allen H. Lerman, and Susan Nelson (2012), Taxpayer Compliance Costs for Corporations and Partnerships: A New Look." In *IRS Research Bulletin: Proceedings of the 2012 IRS Research Conference,* 4–18. Internal Revenue Service, Washington, DC.

Council of Economic Advisers (2013), *Economic Report of the President* (March), 105–107.

Government Accountability Office (GAO) (2005), "Summary of Estimates of the Costs of the Federal Tax System." GAO-05-878: Aug 26, 2005. https://www.gao.gov/products/GAO-05-878.

Guyton, John, John O'Hare, Michael Stavrianos, and Eric Toder (2003), "Estimating the Compliance Cost of the U.S. Individual Income Tax," *National Tax Journal,* Volume LVI, No. 3, September 2003, 673–688.

Lu, Chenxi (2017), "Itemized Deductions." http://www.taxpolicycenter.org/publications/itemized-deductions/full.

Marcuss, Rosemary, George Contos, John Guyton, Patrick Langetieg, Allen Lerman, Susan Nelson, Brenda Schafer, and Melissa Vigil (2013), "Income Taxes and Compliance Costs: How Are They Related?" *National Tax Journal*, 66:4, pp. 833–854

Slemrod, Joel and Marsha Blumenthal (1996), "The Income Tax Compliance Cost of Big Business." *Public Finance Quarterly* 24 (4), 411–438.

Slemrod, Joel, and Varsha Venkatesh (2002), "The Income Tax Compliance Cost of Large and Mid-Size Businesses." Ross School of Business Paper No. 914. University of Michigan, Ann Arbor, MI.

Testa, V., and T. Haines (2012), *Statistics of Income—2012 Individual Income Tax Returns,* Internal Revenue Service, Publication 1304, pp. 23–27.

Appendix

Table A1 shows the different PUF items used in the ITB and the TPC adjusted model, and the corresponding complexity assignments. The items are organized by where they appear in the tax code, starting with the 1040, Schedule A, Schedule C, Schedule D, Schedule E (page 1), Schedule E (page 2), Form 2441 and Form 8863.

Tables A2-A5 show the regression coefficients and variances for the 2007, 2010, 2011, and 2012 ITB survey, before and after the PUF adjustments. Bolded coefficients signify that they are statistically significant, with a t-statistic above 2.00.

TABLE A1. Complexity Categories

Field Description	PUF Field Name	Complexity Assignment
Form 1040		
Salaries, Wages, and Tips	E00200	Low
Interest Received	E00300	Low
Tax-Exempt Interest	E00400	Low
Dividends	E00600	Low
Qualified Dividends	E00650	Medium
State Income Tax Refunds	E00700	Low
Alimony Received	E00800	Medium
Capital Gain Distributions (not reported on Schedule D)	E01100	Low
Supplemental Schedule (Form 4797) Net Gain or Loss	E01200	Medium
Pensions and Annuities—Gross = Net	E01500	Low
Pensions and Annuities—Gross ~= Net	E01500	Medium
Unemployment Compensation	E02300	Low
Social Security Income—Gross = Net	E02400	Low
Social Security Income—Gross ~= Net	E02400	Medium
Total Deductible IRA Payments	E03150	Low
Student Loan Interest Deduction	E03210	Low
Educator Expenses Deduction	E03220	Low
Tuition and Fees Deduction	E03230	Low
Self-Employed Health Insurance Deduction	E03270	Medium
Domestic Production Activity Deduction Amount	E03240	High
Payments to a Keogh Plan and SEP Deduction	E03300	Medium
Forfeited Interest Penalty on Early Withdrawal of Savings	E03400	Low
Alimony Paid	E03500	Medium
Exemption Amount—Amount Not Limited	E04600	Low
Exemption Amount—Amount Limited	E04600	Low
Foreign Tax Credit—Form 1116	E07300	Medium
Gross Child Tax Credit	N24*1000	Low
Credit for Elderly or Disabled—Schedule R	E07200	Medium
Residential Energy Credit—Form 5695	E07260	Medium
General Business Credit—Form 3800	E07400	Medium
Prior Year Minimum Tax Credit—Form 8801	E07600	High
Other Statutory Credits	P08000	Medium
Self-Employment Tax	E09400	Medium

TABLE A1. Complexity Categories—Continued

Field Description	PUF Field Name	Complexity Assignment
Recapture Tax	E09700	Medium
Penalty Tax on Qualified Retirement Plans	E09900	Medium
Income Tax Withheld	E10700	Low
Estimated Tax Payments—Assisted	E10900	Low
Estimated Tax Payments—Unassisted	E10900	Medium
Amount Paid with Form 4868—Request for Extension	E11100	Medium
Excess FICA / RRTA	E11200	Low
Earned Income Tax Credit—Exemption Claimed for Qualifying Child	CODED	Low
Earned Income Tax Credit—Exemption Not Claimed for Qualifying Child	CODED	Medium
Gross Additional Child Tax—Self Employed	N24*1000	Medium
Gross Additional Child Tax—Not Self Employed	N24*1000	Low
Form 4136 Fuels Tax Credit	E11300	Medium
Regulated Investment Company Credit—Form 2439	E11400	Medium
Federal Telephone Excise Tax Credit	E11600	Low
Credit Elect Applied to Next Year's Estimated Tax	E12000	Low
Predetermined Estimated Tax Penalty	E12200	Medium
Form 1040 (Schedule A)		
Total Medical Expenses	E17500	Medium
Income Taxes	E18425	Low
General Sales Taxes	E18450	Medium
Real Estate Taxes	E18500	Low
Cash Contributions	E19800	Medium
Elected Qualified Contributions	E19850	Medium
Other Than Cash Contributions	E20100	Medium
Unreimbursed Employee Business Expense	E20550	Medium
Total Miscellaneous Deductions Subject to 2% AGI Limitation	E20400	Medium
Net Casualty or Theft Loss	E20500	Medium
Form 1040 (Schedule C)		
Income/Loss from Sales and Operations—No COGS	E90040	Low
Income/Loss from Sales and Operations—With COGS	E90040	Medium
Cost of Goods Sold and/or Operations	E90050	Medium
Other Income / (Loss)	E90080	Medium
Depreciation	E90190	Medium
Insurance	E90210	Medium
Mortgage Interest	E90240	Medium
Office Expenses	E90280	Medium
Net Wages	E90370	Medium
Form 1040 (Schedule D)		
Net Short Term Gain or Loss	P22250	Medium
Schedule D Capital Gain Distributions	E22370	Low
Net Long Term Gain or Loss	P23250	Medium

TABLE A1. Complexity Categories—Continued

Field Description	PUF Field Name	Complexity Assignment
Unrecaptured Section 1250 Gain	E24515	High
28% Rate Gain or Loss	E24518	High
Form 1040 (Schedule E—Page 1)		
Total Rents Received	E25350	Medium
Total Royalties Received	E25360	Medium
Mortgage Interest Paid to Financial Institutions	E25370	Medium
Rental Depreciation	E25500	High
Form 1040 (Schedule E—Page 2)		
Total Passive Income	E25360+E27200	Medium
Total Nonpassive Income	E25370+E32800	Medium
Total Passive Loss	E25500+E62100	Medium
Total Nonpassive Loss	E26390+E87520	Medium
Total Income, Estate and Trust	E26390	Medium
Total Loss, Estate and Trust	E26400	Medium
Farm Net Rental Income or Loss	E27200	Medium
Form 2441		
Qualifying Individual Expenses	E32800	Medium
Form 8863		
Tentative Hope Credit Amount	E87520	Low
Lifetime Learning Total Qualified Expenses	E87530	Low

TABLE A2. Regression Coefficients Based on the 2007 ITB Survey

Item	Before PUF Adjustment		After PUF Adjustment	
	Log Model	Variance Model	Log Model	Variance Model
Intercept	-0.1451	1.2477	-0.0951	1.1790
Log Modified Positive Income | Used a Paid Preparer	**0.1904**	-0.0196	**0.1745**	-0.0197
Log Modified Positive Income | Used a Tax Software	**0.2794**	-0.0887	**0.2773**	-0.0873
Log Modified Positive Income | Prepared by Hand	**0.2591**	-0.0151	**0.2486**	-0.0009
Log of Low Complexity Activity	**0.0104**	-0.0038	**0.0113**	-0.0049
Log of Medium Complexity Activity	**0.0077**	0.0010	**0.0140**	0.0016
Log of High Complexity Activity	**0.0207**	0.0043	**0.0234**	-0.0003
Log of Line Count | Used a Paid Preparer	0.0239	0.0313	**0.0614**	0.0395
Log of Line Count | Used Tax Software	**0.1508**	0.0341	**0.1566**	0.0281
Log of Line Count | Prepared by Hand	**0.2996**	-0.0300	**0.3035**	-0.0379
Used a Paid Preparer	**2.7825**	-0.4133	**2.7044**	-0.3399
Used Tax Software	**0.7685**	0.3691	**0.6973**	0.4852
Adj. R-Squared	0.5644	0.0172	0.5575	0.0183

NOTE: Estimates with a t-statistic in excess of 2 are bold.

TABLE A3. Regression Coefficients Based on the 2010 ITB Survey

Item	Before PUF Adjustment		After PUF Adjustment	
	Log Model	Variance Model	Log Model	Variance Model
Intercept	-0.2708	1.5262	0.0241	1.2626
Log Modified Positive Income | Used a Paid Preparer	**0.1455**	-0.0676	**0.1252**	-0.0539
Log Modified Positive Income | Used a Tax Software	**0.2905**	-0.0384	**0.2981**	-0.0335
Log Modified Positive Income | Prepared by Hand	**0.1983**	-0.0905	**0.1588**	-0.0559
Log of Low Complexity Activity	**0.0060**	-0.0012	**0.0071**	-0.0020
Log of Medium Complexity Activity	**0.0081**	0.0006	**0.0135**	0.0016
Log of High Complexity Activity	**0.0275**	0.0084	**0.0340**	0.0022
Log of Line Count | Used a Paid Preparer	**0.2561**	-0.0128	**0.3126**	-0.0014
Log of Line Count | Used Tax Software	**0.2236**	-0.1094	**0.2277**	-0.1125
Log of Line Count | Prepared by Hand	**0.4834**	0.0627	**0.5078**	0.0485
Used a Paid Preparer	**2.2942**	-0.0896	**1.9558**	0.0069
Used Tax Software	0.4410	0.1594	0.0433	0.4142
Adj. R-Squared	0.5430	0.0166	0.5362	0.0149

NOTE: Estimates with a t-statistic in excess of 2 are bold.

TABLE A4. Regression Coefficients Based on the 2011 ITB Survey

Item	Before PUF Adjustment		After PUF Adjustment	
	Log Model	Variance Model	Log Model	Variance Model
Intercept	**-0.9834**	1.1148	**-0.5870**	1.2369
Log Modified Positive Income | Used a Paid Preparer	**0.2371**	-0.0536	**0.2183**	-0.0638
Log Modified Positive Income | Used a Tax Software	**0.4016**	-0.0631	**0.4059**	-0.0679
Log Modified Positive Income | Prepared by Hand	**0.2668**	-0.0330	**0.2105**	-0.0476
Log of Low Complexity Activity	**0.0054**	-0.0031	**0.0063**	-0.0037
Log of Medium Complexity Activity	**0.0089**	0.0012	**0.0148**	0.0017
Log of High Complexity Activity	**0.0240**	0.0107	**0.0464**	0.0146
Log of Line Count | Used a Paid Preparer	**0.1814**	0.0509	**0.2167**	0.0665
Log of Line Count | Used Tax Software	**0.2298**	0.0098	**0.2188**	0.0086
Log of Line Count | Prepared by Hand	**0.4948**	0.0334	**0.5334**	0.0446
Used a Paid Preparer	**2.5055**	-0.1355	**2.1222**	-0.2063
Used Tax Software	0.0257	0.3763	-0.4018	0.3236
Adj. R-Squared	0.5916	0.0221	0.5813	0.0231

NOTE: Estimates with a t-statistics in excess of 2 are bold.

TABLE A5. Regression Coefficients Based on the 2012 ITB Survey

Item	Before PUF Adjustment		After PUF Adjustment	
	Log Model	Variance Model	Log Model	Variance Model
Intercept	-0.5160	1.1975	-0.1582	1.4992
Log Modified Positive Income | Used a Paid Preparer	**0.1765**	-0.0237	**0.1049**	-0.0273
Log Modified Positive Income | Used a Tax Software	**0.3278**	-0.0911	**0.3158**	-0.0877
Log Modified Positive Income | Prepared by Hand	**0.2652**	-0.0325	**0.2204**	-0.0638
Log of Low Complexity Activity	**0.0061**	-0.0017	**0.0078**	-0.0033
Log of Medium Complexity Activity	**0.0108**	0.0013	**0.0182**	0.0017
Log of High Complexity Activity	**0.0267**	0.0066	**0.0581**	0.0105
Log of Line Count | Used a Paid Preparer	**0.1060**	0.0086	**0.1690**	0.0423
Log of Line Count | Used Tax Software	**0.2239**	-0.0331	**0.2157**	-0.0207
Log of Line Count | Prepared by Hand	**0.3748**	0.0110	**0.3875**	0.0394
Used a Paid Preparer	**2.9691**	-0.3897	**3.0045**	-0.7051
Used Tax Software	0.4013	0.7205	0.1297	0.4331
Adj. R-Squared	0.5678	0.0215	0.5405	0.0207

NOTE: Estimates with a t-statistics in excess of 2 are bold.

Counting Elusive Nonfilers Using IRS Rather Than Census Data

Patrick Langetieg, Mark Payne, and Alan Plumley (IRS Research, Applied Analytics, and Statistics)[1]

Although not all individuals have a Federal income tax filing requirement, every year millions of required returns are not filed on time or at all. The Internal Revenue Service (IRS) would like to help taxpayers meet their filing obligations, so estimates of the extent of nonfiling[2] provide helpful insights. For many years, the IRS has used Census Bureau survey data when estimating the extent of individual income tax filing compliance. This paper describes an updated methodology for estimating the number of nonfilers and the corresponding Voluntary Filing Rate using IRS data alone—without relying on Census data. The reasons for this new approach are also explained.

The Voluntary Filing Rate (VFR)

The IRS has estimated the VFR since the mid-1990s to examine factors that influence individual income tax filing compliance.[3] It is defined for a given tax year as:

$$\text{VFR} = \frac{\text{Number of Required Returns Filed on Time}}{\text{Total Number of Returns Required To Be Filed}}$$

We derive the VFR numerator from IRS population data encompassing all filed returns, which we classify as *timely* or not based on a comparison of the filing date and the relevant filing deadline (accounting for all valid extensions) and as *required* or not based on a comparison of all relevant income (reported by the taxpayer) and the filing thresholds in place for the year in question. We have historically estimated the VFR denominator from the Census Bureau's Current Population Survey Annual Social and Economic Supplement (CPS ASEC, hereinafter referred to as CPS), grouping individuals into assumed tax units (e.g., marrieds, singles, or heads of households) and applying comparable logic to estimate whether a tax return was required.

Preliminary estimates for the denominator were first constructed in a necessarily approximate manner since the CPS lacks some of the information needed to confirm various tax-related concepts. Initially, both the numerator and denominator were estimated from samples each year. However, when the IRS began storing data on the whole population in a form that was accessible to IRS research staff, we began estimating the numerator from population data. That allowed us to examine in more detail what type(s) of taxpayers were driving fluctuations in the numerator.

Subsequent work addressed various measurement issues surrounding the numerator and denominator of the VFR ratio. This work included ensuring that the numerator and denominator more precisely represented the same population of taxpayers (U.S. residents over the age of 14), and that they reflected the same definitions (as much as the data would allow) for the requirement to file and what it means for a required return to be timely filed for VFR purposes.[4]

---

[1] The views expressed in this paper are those of the authors, and do not necessarily represent the positions of the Internal Revenue Service. This paper is an extension of our earlier papers: Erard, et al. (2012), and Erard, et al. (2014).

[2] We use the term "nonfiler" to include only those who are required to file a Form 1040 for income tax or employment tax purposes, but did not file a return on time. Therefore, nonfilers include taxpayers who filed a required return after the due date (plus extension). This compliance-oriented definition of nonfilers differs from a policy-oriented definition, which includes those who might be eligible for (current or proposed) benefits offered through the tax system without incurring a tax obligation. Timely filers exclude those who have no filing requirement, but who file solely to claim a refund of withholding or to claim a refundable credit.

[3] See Plumley (1996).

[4] We include in the numerator only those required returns that are filed on time by December 31 of the primary filing year. This excludes returns that are considered timely (e.g., due to combat extensions), but are filed much later than most, and also excludes returns filed by December 31 but after their filing deadline. Setting December 31 as the cut off date allows for a consistent measure to be produced each year.

Accounting for Missing Income

Even after the refinements mentioned above, it became necessary to augment the Census data used to construct the denominator of the VFR to account more fully for certain types of income that are reported to the IRS, but not fully reported on Census surveys (such as pensions, Social Security income, sole proprietor income, and unemployment compensation). An understatement of income in the denominator of the measure would contribute to an overstatement of the VFR. Erard, *et al.* (2014) documents the differences in the amounts of these types of income reported on the CPS survey versus what is reported on the third-party information returns sent to the IRS (in the case of pension and Social Security income, as well as unemployment compensation), or what is reported on filed income tax returns (in the case of self-employment income). To address these discrepancies, we developed econometric methodologies for imputing the missing income to the CPS records.

Insights from Tax Gap Research

In addition to estimating the number of nonfilers, the IRS also estimates the nonfiling tax gap—the estimated amount of tax not paid on time by those who do not file on time. The most recent estimate of the individual income tax nonfiling gap was calculated two ways: one based primarily on Census data matched to IRS data,[5] and another based mostly on IRS administrative data (such as income reported by third parties).[6] A curious observation from those estimates was that while they yielded very similar estimates of both the number of nonfilers and the corresponding dollar value of the gap, they suggested that the number of nonfilers was much larger than had been previously understood and the Census Method suggested that the number of required returns was much less than the number derived from the Administrative Data Method. That caused us to compare estimates of the VFR numerator and denominator derived in three ways: (1) using the latest VFR methodology; (2) using the Census Method for estimating the nonfiling gap; and (3) using the Administrative Data Method for estimating the gap. As Table 1 shows, both tax gap methods yielded an estimated 15 million nonfilers for Tax Year 2010, while the latest VFR methodology estimated only 6.3 million nonfilers for the same year. However, the Census Method for estimating the tax gap suggested that both numerator and denominator were around 11 million returns fewer than derived from the Administrative Data Method. This undercounting of required returns using Census data affected the denominator of the VFR, but not the numerator (which is tabulated from IRS data), resulting in an inflated VFR relative to either of the tax gap methods.[7] This observation exposes the inadequacy of the Census data to estimate the number of required returns—even after imputing pension, Social Security, sole proprietor, and unemployment compensation income to the Census data. Given the total counts observed in IRS data, it seems clear that many individuals who appear in the adjusted CPS data as *not* having enough income to be required to file actually *do* have a filing requirement. This affects both the number of required returns among those who file on time (the numerator of the VFR) and the estimated number of nonfilers. The common thread is the underreporting of income in the Census data.

TABLE 1. Thousands of Returns in VFR Components Estimated by Different Methods, Tax Year 2010

Item	VFR Method	Census Method	Administrative Data Method
Numerator (required returns filed on time)	115,900	105,001	115,900
Denominator (total required returns)	122,200	119,967	130,787
Difference (implied number of nonfilers)	6,300	14,966	14,937
Numerator/Denominator (Implied VFR)	94.8%	87.5%	88.6%

[5] The Census Bureau has developed the ability to assign an anonymous Protected Identification Key (PIK) to most respondents in the CPS and to all of the records Census routinely receives from the IRS for the population (including selected data from income tax returns and from third-party information documents). This allows Census to create a matched anonymized dataset containing both Census and IRS data for a representative sample of the population. However, there are some CPS records that cannot be matched to the IRS data because a PIK could not be assigned to them with adequate certainty. We therefore restricted our analysis to the records that could be matched, and re-weighted them to represent the entire population.

[6] See Langetieg, et al. (2016).

[7] The VFR method for estimating the denominator is very similar to the tax gap Census Method. The main difference is that the VFR method uses the entire CPS dataset, while the Census Method uses a subset of the CPS that is matched to IRS return data, which makes it possible to impute income to the Census records in a more sophisticated way.

Attempts To Impute More Income to the Census Data

Several attempts to impute still more income to the CPS yielded unsatisfactory results. One approach was to use the matched subsample, basing total income on the amount reported on the tax return (when available) instead of what was reported in the Census survey. Another approach was a backend imputation of gross income (without regard to source) that was calibrated to totals in IRS data. Neither produced a time series of required returns that was sufficiently similar to that produced from IRS data alone, particularly for key years when the number of required filers was known to have spiked.[8] Although further research may allow us to make adequate adjustments to the CPS data that would yield accurate estimates of the VFR, we concluded that it was necessary to estimate the VFR based solely on IRS administrative data.

Table 2A provides the population estimates based on the matched sub-sample. The numerator, denominator, and implied VFR for several approaches are provided, including the two attempts to correct (adjust) gross income mentioned above. It's important to note that the numerator and denominator estimates provided in Table 2A are from Census survey data so the numerator is not the actual IRS estimate.

TABLE 2A. VFRs and Millions of Returns in VFR Components Estimated by Census-Based Methods, Tax Years 2007 to 2014

Item	Tax Year							
	2007	2008	2009	2010	2011	2012	2013	2014
All Census Tax Units	239.0	240.6	243.9	245.9	247.8	249.8	251.4	253.9
All Census Tax Units that Filed a Return	146.0	135.4	133.7	133.8	137.0	138.2	138.2	140.6
	Previous Census-Based Methodology							
Numerator (Filed Required Returns)	113.5	111.5	108.7	108.9	110.9	112.2	112.7	114.9
Denominator (Total Required Returns)	124.0	124.2	121.0	120.9	123.2	124.8	126.0	129.8
VFR	91.5%	89.8%	89.8%	90.1%	90.0%	89.9%	89.4%	88.5%
	Updated Census-Based Methodology							
Numerator (Filed Required Returns)	114.3	112.5	109.8	110.2	112.1	113.4	114.0	116.4
Denominator (Total Required Returns)	125.9	127.0	124.3	124.7	126.6	128.5	130.0	133.7
VFR	90.8%	88.6%	88.3%	88.4%	88.5%	88.2%	87.7%	87.1%
	Updated Census Based-Methodology Form 1040 Income Amount Used When Available							
Numerator (Filed Required Returns)	125.1	118.5	116.7	117.8	120.0	120.8	120.5	123.6
Denominator (Total Required Returns)	136.7	133.0	131.2	132.3	134.5	135.9	136.4	140.9
VFR	91.5%	89.1%	88.9%	89.0%	89.2%	88.9%	88.3%	87.7%
	Updated Census-Based Methodology Backend Adjustment to Gross Income to All Tax Units							
Numerator (Filed Required Returns)	117.2	113.8	111.2	111.8	113.9	115.2	115.6	118.2
Denominator (Total Required Returns)	137.8	134.1	131.7	132.7	134.9	136.6	137.8	141.8
VFR	85.1%	84.9%	84.4%	84.3%	84.4%	84.3%	83.9%	83.4%
	Updated Census-Based Methodology Form 1040 Income Amount Used When Available Backend Adjustment to Gross Income to Tax Units Without a 1040							
Numerator (Filed Required Returns)	126.5	119.4	117.4	118.4	120.8	121.7	121.4	124.4
Denominator (Total Required Returns)	147.0	139.8	137.9	139.3	141.8	143.1	143.6	148.1
VFR	86.1%	85.4%	85.1%	85.0%	85.2%	85.0%	84.5%	84.0%

[8] For example, the number of required filers spiked in Tax Year 2007 because in order to receive the one-time Economic Stimulus Payment, people had to file a tax return for 2007. Many more returns were filed that year that had a requirement to file for regular tax reasons so the VFR spiked that year then fell somewhat in the subsequent years.

The Previous Methodology estimates for 2010 are slightly different from the estimates provided in Table 1 for two reasons. First, the Table 2A numerators include all required returns filed during the calendar year they were due so both timely and late required returns are included in the numerator. Second, the reweighting methodology to account for missing PIKs[9] was changed to an inverse probability method for this round of estimates.

The difference between the Previous and Updated Methodology estimates is due to a change to the sole proprietor income imputation. Previously, the sole proprietor imputation for the Census Method was calibrated to match the proportion of individuals with sole proprietor income observed on Form 1040. This approach was not used for the Administrative Data Method. In order to be consistent across methods, the Census Method was changed to match the Administrative Data Method. This change leads to a larger number of individuals receiving imputed sole proprietor income and, as a result, more required returns.

The Previous Method and Updated Method estimates show the same denominator deficiency discussed from Table 1. The last three sets of results show the effect of using Form 1040 income when available and/or computing a backend adjustment to gross income. These methods appear to resolve the denominator deficiency, some more reasonably than others.

Table 2B provides the same information shown in Table 2A, except the Census-based numerator estimates have been replaced with the actual IRS estimates.

TABLE 2B. VFRs and Millions of Returns in VFR Denominators Estimated by Census-Based Methods and Numerators Estimated From IRS Data, Tax Years 2007 to 2014

Item	Tax Year							
	2007	2008	2009	2010	2011	2012	2013	2014
All Tax Units	239.0	240.6	243.9	245.9	247.8	249.8	251.4	253.9
Tax Units that Filed a Return	146.0	135.4	133.7	133.8	137.0	138.2	138.2	140.6
	Previous Census Based-Methodology							
IRS Numerator (Filed Required Returns)	117.2	116.3	113.4	115.4	117.2	117.4	118.5	120.5
Denominator (Total Required Returns)	124.0	124.2	121.0	120.9	123.2	124.8	126.0	129.8
VFR	94.5%	93.6%	93.7%	95.5%	95.1%	94.1%	94.0%	92.8%
	Updated Census-Based Methodology							
IRS Numerator (Filed Required Returns)	117.2	116.3	113.4	115.4	117.2	117.4	118.5	120.5
Denominator (Total Required Returns)	125.9	127.0	124.3	124.7	126.6	128.5	130.0	133.7
VFR	93.1%	91.6%	91.2%	92.5%	92.6%	91.4%	91.2%	90.1%
	Updated Census-Based Methodology Form 1040 Income Amount Used When Available							
IRS Numerator (Filed Required Returns)	117.2	116.3	113.4	115.4	117.2	117.4	118.5	120.5
Denominator (Total Required Returns)	136.7	133.0	131.2	132.3	134.5	135.9	136.4	140.9
VFR	85.7%	87.4%	86.4%	87.2%	87.1%	86.4%	86.9%	85.5%
	Updated Census-Based Methodology Backend Adjustment to Gross Income to All Tax Units							
IRS Numerator (Filed Required Returns)	117.2	116.3	113.4	115.4	117.2	117.4	118.5	120.5
Denominator (Total Required Returns)	137.8	134.1	131.7	132.7	134.9	136.6	137.8	141.8
VFR	85.1%	86.7%	86.1%	87.0%	86.9%	85.9%	86.0%	85.0%
	Updated Census-Based Methodology Form 1040 Income Amount Used When Available Backend Adjustment to Gross Income to Tax Units Without a 1040							
IRS Numerator (Filed Required Returns)	117.2	116.3	113.4	115.4	117.2	117.4	118.5	120.5
Denominator (Total Required Returns)	147.0	139.8	137.9	139.3	141.8	143.1	143.6	148.1
VFR	79.7%	83.2%	82.2%	82.8%	82.7%	82.0%	82.5%	81.4%

[9] Refer to footnote 5.

The Table 2B estimates show that each of the attempts to account for the denominator deficiency led to a lower overall VFR in 2007. As discussed later, there is good reason to expect the 2007 VFR to be higher in 2007 than the following years, so this outcome suggests that the Census Method does not accurately estimate the number of required returns in the population (the VFR denominator). Further research is needed to understand and fully address this deficiency before moving forward with the Census Method.

The Administrative VFR Methodology

The VFR based on IRS administrative data relies on information from filed individual income tax returns and third-party information returns. The population is made up of filed returns, whether timely or late, and those who do not appear on a filed return—"not-filers." As in the CPS-based estimate, the numerator is the count of all required returns that were filed on time (including extensions). The denominator is the count of all required returns in the population.[10] For filed returns, the determination of whether the return was required is based on the income reported on the return. For most filers, this involves comparing the amount of gross income with thresholds linked to filing status and whether the taxpayer is over 65 years of age. Taxpayers with self-employment income are required to file if their net self-employment income is $433 or greater.

For individuals who do not appear on a filed return, third-party information documents are the main sources for estimating income. But, given the fact that most self-employment income is not reported on information documents, this type of income is imputed to not-filers using models developed from filed returns. In addition, filing status and family units are imputed. The method for determining the number of required returns among not-filers involves the following steps:

- Identify all individuals who appeared on a third-party information return for the tax year in question, but who did not appear on a filed return as a primary or secondary taxpayer for the given tax year by the end of the second year after the conclusion of the tax year.[11] This excludes (as potential primary taxpayers or spouses) those for whom no third-party information return was filed for the year in question.[12] The two-year cutoff for using information on late returns is put in place to limit the lag time after the conclusion of filing before the VFR estimate can be made while also ensuring consistency across the time series. The disadvantage is that the potentially more accurate information on income and tax unit structure available on the returns that are filed after this two-year cutoff is sacrificed.
- Identify the known income, prepayments, and State of residence for these not-filers from third-party information documents and other administrative tax data sources (e.g., Master File). In addition, the social security master file (DM-1) was used to obtain the age and gender of each individual. Finally, the individuals were matched to filed returns to determine which ones had been claimed as dependents.
- Impute self-employment income to the individual potential not-filers.
- Assign spouses and filing status and a number of dependents to the remaining not-filers using their age and gender, so that the combined age and filing status distribution of timely filers, late filers, and not-filers matches the corresponding distributions of singles, marrieds, heads of households, and dependents in CPS data, after the aggregate counts from those on filed returns is taken out.
- Convert net sole proprietor and farm income below the $433 threshold to gross income using multipliers based on mean ratios of gross to net observed on filed returns.
- Determine whether a tax return was required to be filed by comparing gross income for the tax unit to the filing thresholds for the relevant filing status and checking whether the net self-employment income threshold was exceeded.

The combined count of primary and secondary taxpayers on filed returns, dependents claimed on filed returns who do not file on their own, and not-filers who are not claimed as dependents is fairly close to

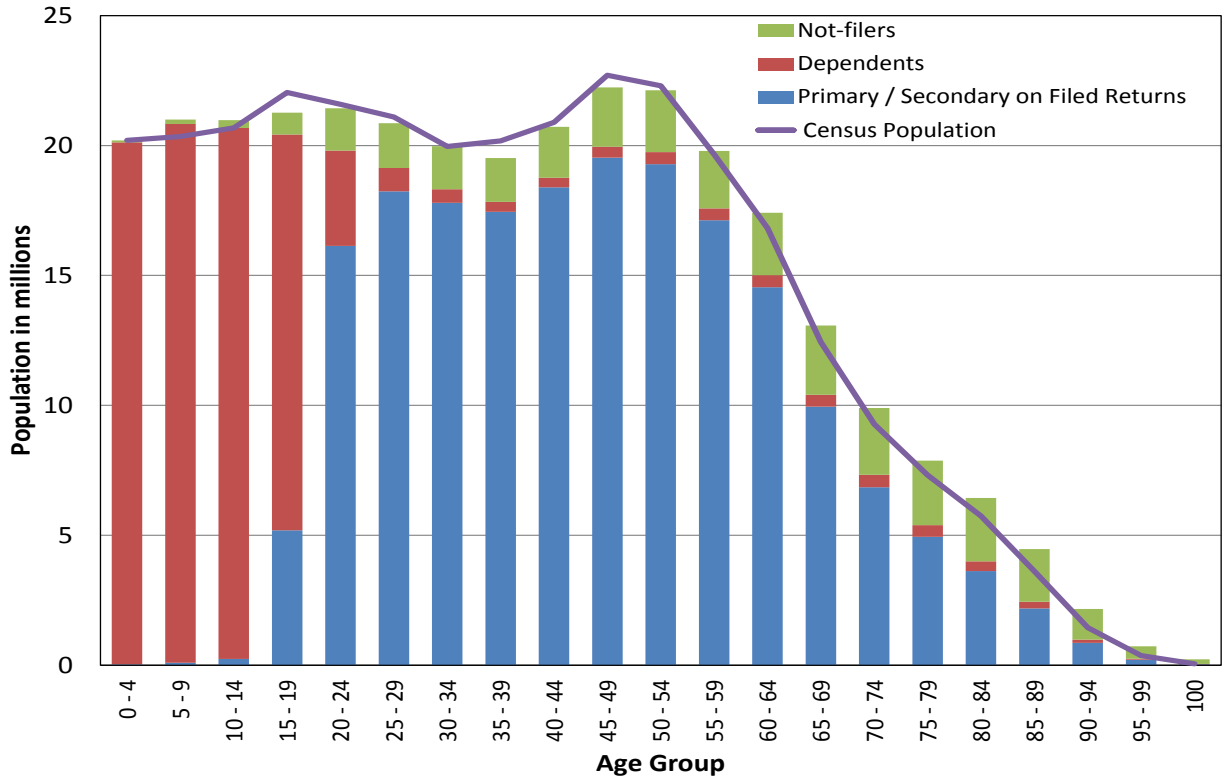
[10] Returns with foreign addresses are excluded from the counts of required filers in the numerator and denominator.

[11] Note that this treats all dependents as potentially required to file a return in their own right.

[12] This approach excludes people who had income only from sources not subject to third-party reporting (such as self-employment income).

Census-based estimates of U.S. population totals at most age levels (see Figure 1). Even though some not-filers are likely not present on information returns, their numbers appear to be offset by other individuals who are present on information returns but not counted in the U.S. Census, such as those who are living overseas. Hence, the population included in the VFR methodology is close to that of the whole U.S. population.

FIGURE 1. Population by Age: IRS Administrative Data vs. Decennial Census, TY2010



Imputation of Net Self-Employment Income

Most self-employment income is not reported on third-party information documents. About 12.9 million of the 25.7 million individuals reporting Schedule C income received a Form 1099 Miscellaneous Income in TY 2014 and an additional 800,000 were matched to a Form 1099 K. Approximately 400,000 were matched to both. However, these forms often provide just a partial picture of the income earned. And, in some cases the forms overestimate net income because they do not provide information about expenses and other deductions to gross income. Given the deficient coverage of information documents, we impute self-employment income to the not-filers based on models derived from filed tax returns.

The limitations of this approach include the fact that self-employment earnings are significantly under-reported on tax returns and that not-filers are likely different from filers. On the one hand, one reason why taxpayers may not file is because their net income is low enough to keep them below the filing threshold. At the same time, not-filers may be more likely to be self-employed and to have most of their income not subject to third-party reporting since this is the characteristic that provides the opportunity to avoid detection by IRS nonfiler enforcement programs. Thus, while our imputations aim to make the sole proprietor earnings of not-filers consistent with what is observed in tax return data, this measure may fall short of true earnings, meaning that we are not fully able to account for all returns that have a filing requirement solely as a result of the $433 self-employment earnings threshold.

The econometric framework for imputing self-employment income involves prediction equations for the presence and magnitude of this form of income as a function of individual characteristics, such as age, gender, and region, as well as sources and amounts of earnings, including wages, interest, Social Security and pension

income, unemployment compensation, and nonemployee compensation. In addition, the imputation of self-employment income is tailored to the ways in which this type of income relates to the filing requirement. If net sole proprietor income exceeds $433, that by itself creates a filing requirement because some self-employment tax is owed. Sole proprietors are also required to file if gross income from self-employment and other sources exceeds the filing threshold for their filing status. Thus, we also convert the imputed net amounts of sole proprietor income into gross income amounts by multiplying them by the average ratio of gross to net income in the population individual return data. Different gross-up factors are applied to different net self-employment earnings categories, with negative factors being applied when net self-employment income is negative (thereby resulting in a positive value for gross self-employment income) and a positive factor being applied when net self-employment income is between $0 and $433.[13] Thus, the econometric framework aims to estimate the likelihood that an individual has earnings falling into one of the following three categories: negative net self-employment earnings, net self-employment earnings between $1 and $433, and net self-employment earnings in excess of $433.

The econometric framework involves three separate models. The first is a probit specification for the likelihood that a filing unit has nonzero self-employment earnings:

$$SE^{*} = \gamma' x + \mu \tag{1}$$

where SE^{*} is a latent variable describing the propensity for net self-employment earnings to be present, x is a vector of explanatory variables, and γ is a vector of coefficients to be estimated. The error term u is assumed to follow the standard normal distribution. Estimation of this model permits us to develop a prediction equation for the unconditional likelihood that an individual has net income from self-employment. Results for this model applied to Tax Year 2014 data are shown in Table 3.

TABLE 3. Probit Model for the Presence of Self-Employment Income, Tax Year 2014

Variable	Parameter Estimate	Chi-square
Intercept	-1.475	(409.2)**
logage	0.152	(63.5)**
male	-1.101	(149.4)**
logage*male	0.341	(210.5)**
west	-0.066	(33.8)**
midwest	-0.108	(84.8)**
northeast	-0.042	(11.7)**
wagesind	0.585	(185.8)**
interestind	0.195	(68.1)**
socsecind	-0.471	(7.6)**
pensionind	0.215	(19.4)**
F1099miscind	0.717	(1847.1)**
unempcompind	0.042	(0.1)
logwages	-0.135	(1045.3)**
loginterest	-0.047	(90.4)**
logpension	-0.044	(70.6)**
logsocsec	-0.028	(2.5)
logunempcomp	-0.011	(0.3)
lognonempcomp	0.148	(4701.8)**
Number of observations	198,704	
R^2	0.217	
Max-rescaled R^2	0.408	

*significant at .05 level; ** significant at .01 level.
See the appendix for variable descriptions.

[13] Different multiplier factors are applied for net self-employment income less than -$5,000 and in the ranges of -$5,000 to -$3,000, -$3,000 to -$1,000, -$1,000 to 0, and $0 to $433.

Estimation of Net Self-Employment Income Category

Our second model is an ordered probit specification for the dollar amount category that net self-employment earnings fall into when they are present: negative, $1 to $433, or over $433:

$$I^*_{SE} = \delta' x + \nu \tag{2}$$

where I^*_{SE} is a latent variable for the propensity for net self-employment earnings to fall into one of these categories, x is the same set of explanatory variables used in our probit model, δ is a coefficient vector to be estimated, and v is a standard normal random disturbance. The model also includes a limit parameter l to be estimated.[14] The indicator I_{SE} for the net self-employment earnings category is assigned as follows:

$$I_{SE} = \begin{cases} 1 & \textit{net earnings} < \$0 \\ 2 & \$0 < \textit{net earnings} <= \$433 \\ 3 & \textit{net earnings} > \$433. \end{cases} \tag{3}$$

The estimation results for Tax Year 2014 are shown in Table 4.

TABLE 4. Ordered Probit Models for the Category of Self-Employment Income, Tax Year 2014

Variable	Parameter Estimate	t-statistic
Intercept	2.187	(12.5)**
logage	-0.346	(-7.6)**
male	-0.742	(-3.4)**
logage*male	0.205	(3.6)**
west	0.037	(1.6)**
midwest	0.122	(5.0)
northeast	0.223	(8.5)**
wagesind	1.384	(16.4)**
interestind	-0.019	(-0.4)
socsecind	-0.085	(-0.3)
pensionind	-0.036	(-0.4)
F1099miscind	-0.072	(-2.2)*
unempcompind	0.339	(1.1)
logwages	-0.212	(-25.8)**
loginterest	0.020	(2.0)*
logpension	-0.026	(-2.4)*
logsocsec	-0.027	(-0.8)
logunempcomp	-0.054	(-1.4)
lognonempcomp	0.082	(21.9)**
Limit	0.207	(38.9)**
Values of dependent variable I_{SE}	Number of observations	
1 = (SE Income < 0)	5,730	
2 = (0 < SE Income <= 433)	1,408	
3 = (SE Income > 433)	17,232	
Total number of observations	24,370	
Missing values	5	
Log Likelihood	-16,146	

* significant at .05 level; ** significant at .01 level.
See the appendix for variable descriptions.

[14] This parameter serves as a threshold for separating the various levels of the response variable.

Imputation of Net Self-Employment Income Amount

Our third model is a regression specification for the magnitude of net self-employment earnings when they exceed $433. Although a taxpayer is required to file a tax return as long as net earnings from self-employment exceed $433, it is desirable to predict their actual magnitude for the tax gap analysis and other research efforts. For instance, this will facilitate an econometric analysis of reporting compliance among self-employed taxpayers. Our specification is:

$$ln(SE) = \beta' x + \varepsilon, \tag{4}$$

where *ln(SE)* represents the natural log of net self-employment earnings, x is the same set of explanatory variables used in the preceding models, β is a vector of coefficients to be estimated, and ε is assumed to be a normal random error term with mean zero and standard deviation σ. Under this specification, the distribution of self-employment earnings is assumed to be log normal. The estimation results for this model for Tax Year 2014 are shown in Table 5.

TABLE 5. Regression Analysis for the Amount of Self-Employment Income > $433, Tax Year 2014

Variable	Parameter Estimate	t-statistic
Intercept	7.723	(44.1)**
logage	0.368	(8.0)**
male	-1.031	(-4.7)**
logage*male	0.340	(5.9)**
west	0.149	(6.3)*
midwest	-0.004	(-0.2)
northeast	0.083	(3.2)**
wagesind	0.303	(3.7)**
interestind	0.127	(2.6)*
socsecind	-2.700	(-7.4)**
pensionind	-0.151	(-1.3)
F1099miscind	-1.171	(-29.3)**
unempcompind	0.228	(0.7)
logwages	-0.110	(-13.2)**
loginterest	0.057	(5.6)**
logpension	0.007	(0.6)
logsocsec	0.201	(5.3)**
logunempcomp	-0.061	(-1.4)
lognonempcomp	0.138	(33.7)**
R^2	0.213	
Adjusted R^2	0.213	
Root MSE	1.213	
Coefficient of Variation	13.515	
N	17,856	

* significant at .05 level; ** significant at .01 level.
See the appendix for variable descriptions.

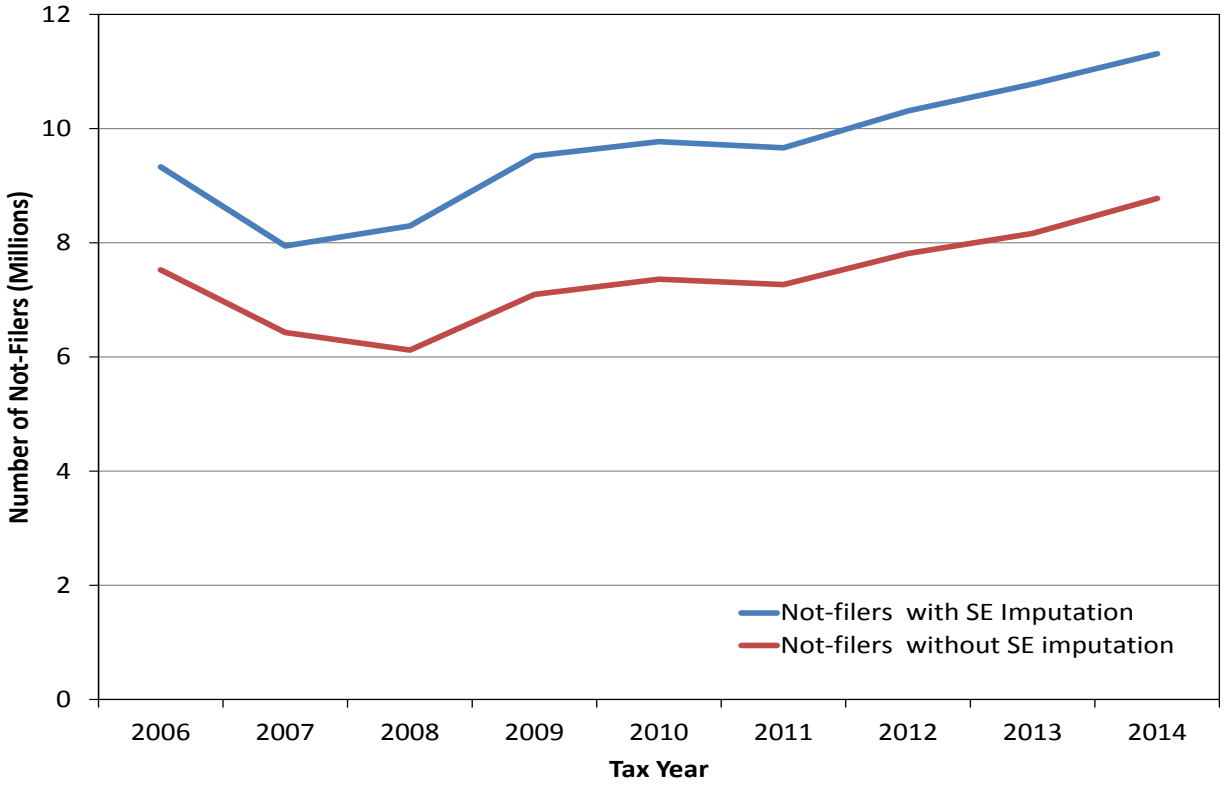
For not-filers flagged for imputation of net self-employment earnings in excess of $433, the parameter estimates are used to impute earnings (SE) as follows:

$$SE = exp\left(\tilde{\beta}' x + e\right), \tag{5}$$

where $\tilde{\beta}'$ is the estimated coefficient vector and e is a random draw from a normal distribution with mean zero and standard deviation equal to the root mean-squared error of the regression.

For not-filers flagged for imputation of net self-employment losses, we assign a random draw from a lognormal distribution with parameters selected based on summary statistics for reported losses from administrative data. Finally, for not-filers flagged for imputation of net self-employment earnings between $1 and $433, we assign the mean reported earnings among taxpayers reporting earnings in that range. Imputing self-employment income causes the number of estimated required not-filer returns to increase by between 25 percent and 35 percent, or about 2.5 million returns, as illustrated in Figure 2.

FIGURE 2. Counts of Not-Filers With and Without Imputation of Self-Employment Income, TYs 2006–2014

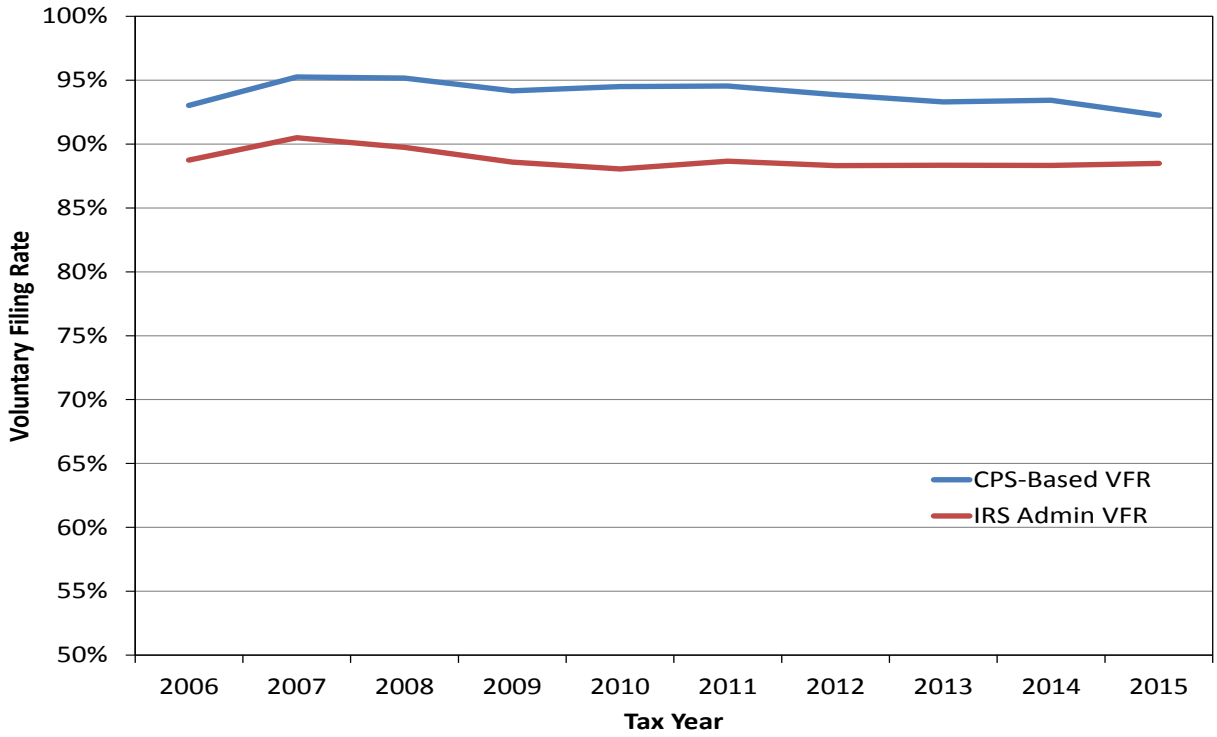


The Resulting VFR

Figure 3 shows the resulting VFR using the IRS Administrative Data Method and compares it with the prior estimates that used IRS filing data for the numerator and CPS as the starting point for the denominator. Given the higher estimate of required returns in the population, the VFR based on IRS data is about 5 percentage points lower than the VFR using CPS data. Both estimates show a significant increase in the VFR in 2007 due to the tax rebates associated with the Economic Stimulus Act of 2008.[15] The series based on IRS data shows a more rapid decline followed by considerable stability. By contrast, the series using CPS shows a slower decline from the peak in Tax Year 2007 with additional reductions more recently. But, by TY 2014, the average difference between the two estimates returned to within 0.3 percent of the difference in TY 2007.

[15] This law provided eligible taxpayers with a rebate of up to $300 per person ($600 for married couples filing jointly) and $300 per dependent child under the age of 17. In order to receive these payments, taxpayers had to file a 2007 tax return.

FIGURE 3. Individual Income Tax Voluntary Filing Rate, TYs 2006–2015



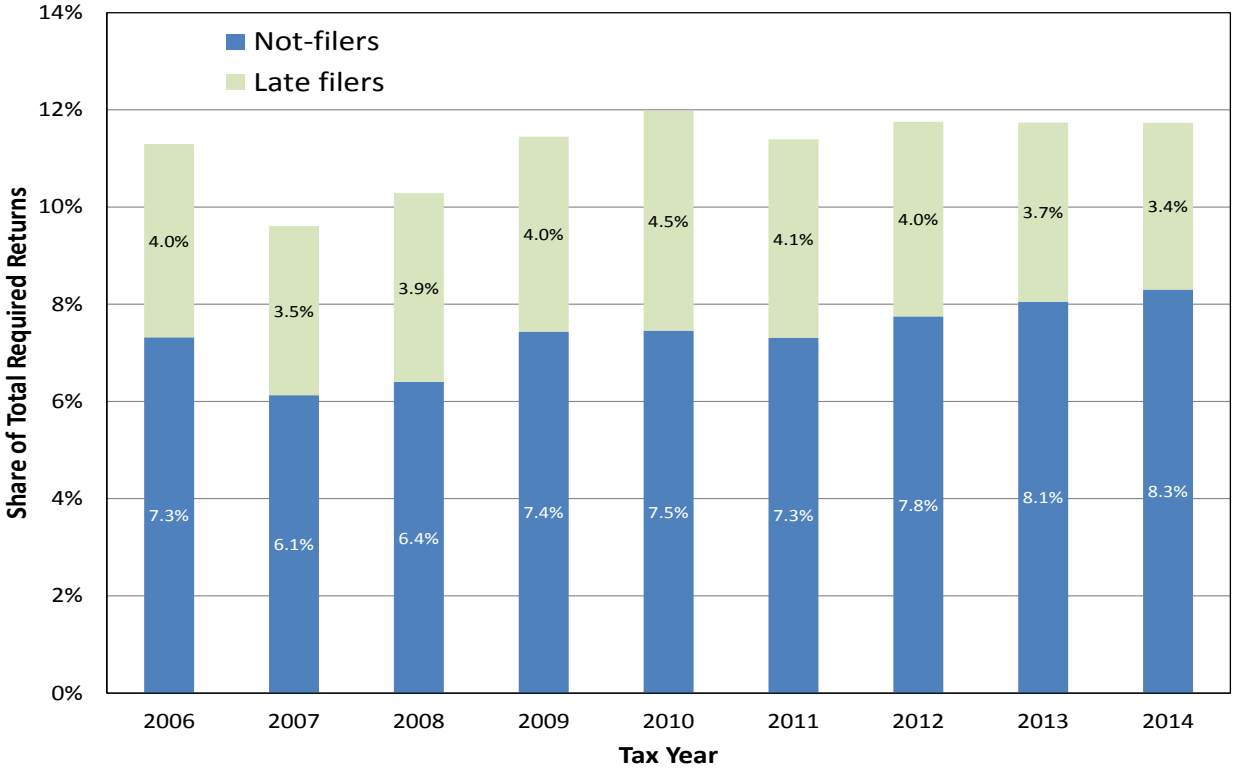
NOTE: The IRS Administrative VFR for Tax Year 2015 is provisional since it does not reflect a full two years of late-filer data.

TABLE 6. Voluntary Filing Rate and Related Estimates, TYs 2006–2015

Tax Year	Millions of Required Returns			VFR (Ratio)
	Total Population (Denominator)	Timely Filed (Numerator)	Nonfilers (Difference)	
2006	127.3	113.0	14.3	88.8%
2007	129.5	117.2	12.3	90.5%
2008	129.4	116.3	13.2	89.7%
2009	128.0	113.4	14.6	88.6%
2010	131.0	115.4	15.6	88.1%
2011	132.1	117.2	15.0	88.7%
2012	132.9	117.4	15.5	88.3%
2013	134.1	118.5	15.6	88.3%
2014	136.4	120.5	15.9	88.3%
2015	138.4	122.5	15.9	88.5%

Given that the VFR has been quite stable over the past four years, this means that the percent of nonfilers has also been equally stable. But, in recent years, the share of nonfilers that remain as not-filers (rather than filing late returns) by two years after the end of the tax year has increased (Figure 4). This may be due to the reduction in IRS resources devoted to nonfiler enforcement. Since many of the returns that are secured from enforcement do not come in until the third year following the tax year, this effect would likely be even more evident if these returns were also included in the late-filing population.

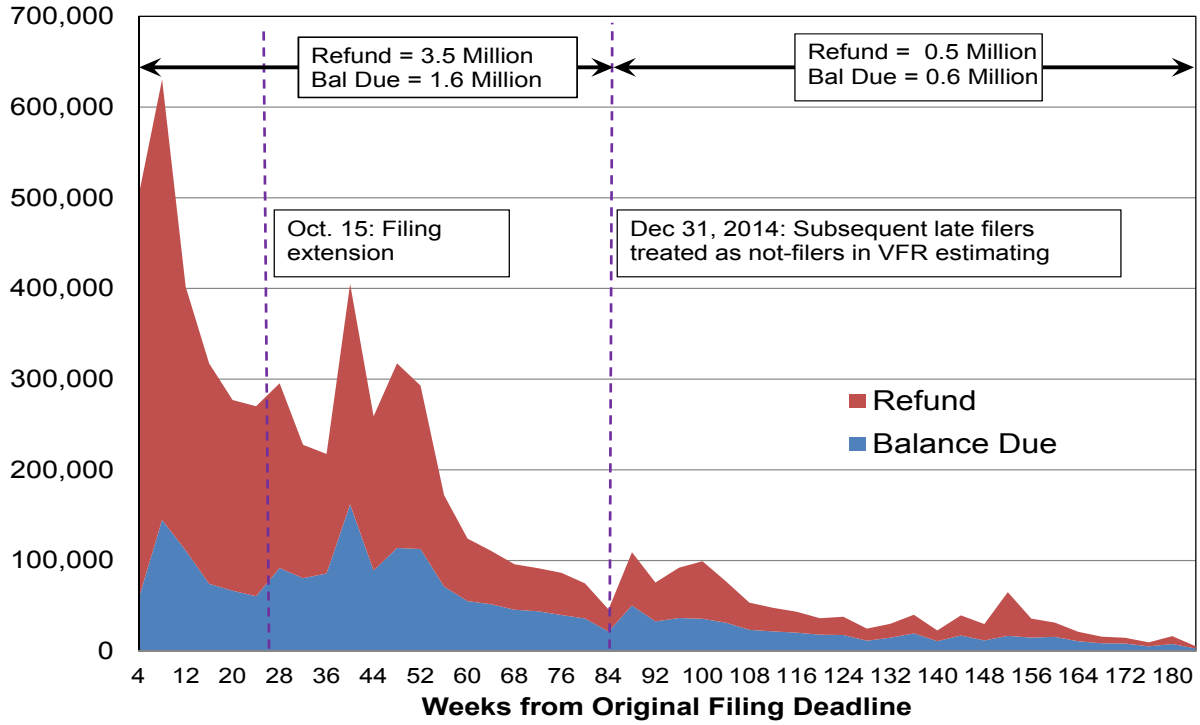
FIGURE 4. Nonfilers as a Percent of Total Required Population, TYs 2006–2014



Figures 5 and 6 show the counts of the late returns over the four-year period following Tax Year 2012. As you can see, in the VFR methodology about 5.1 million of the late returns (those filed within two years of the end of the tax year) are used in the estimation, while the remaining 1.1 million late returns filed three and four years after the end of the tax year are treated as not-filers. Figure 5 shows the breakdown of late returns between refund and balance-due conditions. Refunds are much more common among late returns filed within one year of the end of the tax year (2.2 million out of 2.9 million) than those filed in the second year (1.3 million out of 2.3 million). Refund returns make up only 60 percent of the returns filed in the third and fourth years following the end of the tax year. Figure 6 shows the breakdown between the late returns secured through enforcement efforts and those that come in without enforcement.[16] Because nonfiler notices do not start going out until the end of the filing year, very few returns that are filed within the first year can be attributed to enforcement. But, in the second year about 400,000 out of 2.3 million were filed after the taxpayers were sent nonfiler notices. And, in the third year, about 170,000 out of 700,000 returns had been sent nonfiler notices.

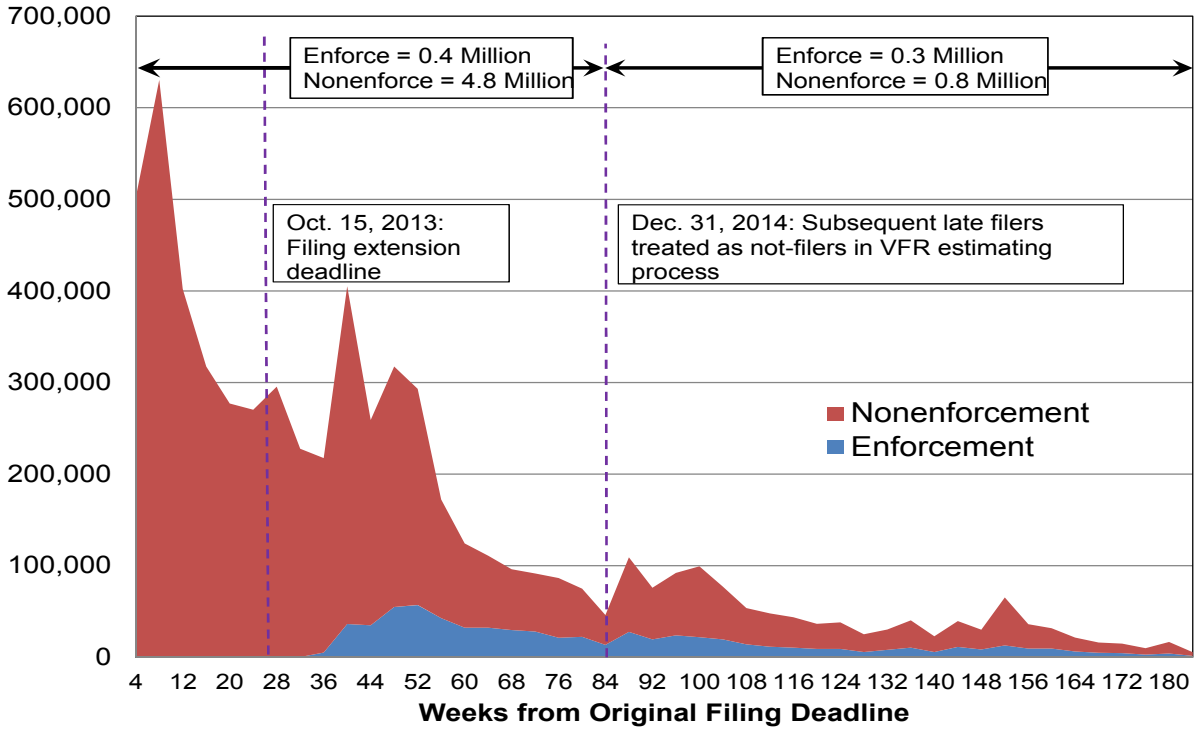
[16] The data shown here are an underestimate of the numbers of returns brought in by enforcement since only those returns that are processed and entered into the Individual Returns Transaction File (IRTF) are included. A fairly large number of returns secured through the Automatic Substitute for Return Program (ASFR) do not get processed in the same way and do not get posted in the IRTF. In addition, a significant number of returns might be filed because of enforcement related to other tax years, but no nonfiler notice was sent for the particular tax year in question.

FIGURE 5. Counts of Balance Due and Refund Returns by Weeks After the Filing Deadline, TY 2012



SOURCE: IRS Individual Returns Transaction File.

FIGURE 6. Counts of Enforcement and Nonenforcement Returns by Weeks After the Filing Deadline, TY 2012



SOURCE: IRS Enforcement Revenue Information System

Exploring Some Drivers of Filing Behavior

Aside from providing a more accurate estimate of the VFR and of the number of required returns in the population, the IRS Administrative Data Methodology offers the additional advantage of using the same data source for both the numerator and denominator of the ratio. This makes it easier to probe the potential causes of fluctuations in the VFR and to examine drivers of filing behavior. For example, if we want to explore the impact of the stimulus payments on the VFR increase in TY 2007, we can examine how many taxpayers who look like they were required to file in TY 2006 but did not file timely ended up timely filing for TY 2007 but then returned to not filing for TY 2008. Then we can compare the net fluctuations of taxpayers in and out of the numerator and denominator in TY 2007 with what we observe for prior and subsequent years.

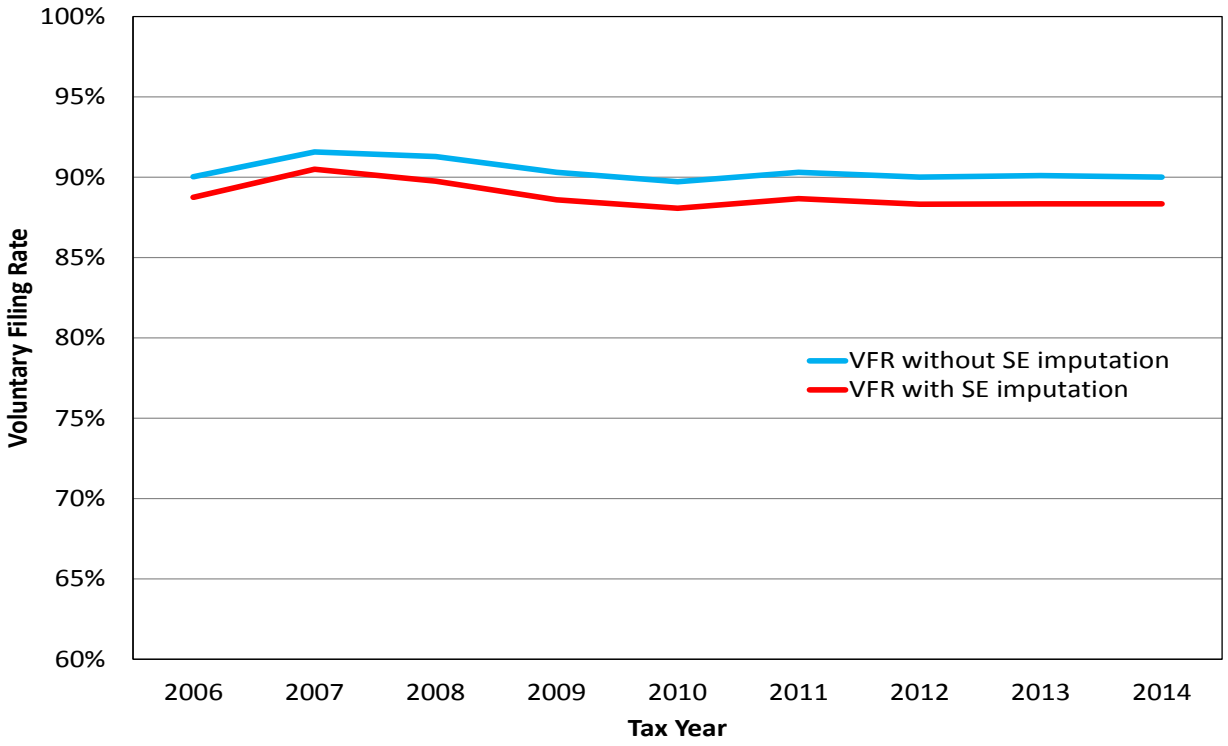
Nonetheless, the new methodology is not without some limitations. First, we do not know the filing status or family make-up—including the identity of the spouse and the number of dependents—of those who do not file. In the current estimates, we impute family units so that in the aggregate the profile of filers and nonfilers matches that evident in Census data. While this method should provide reasonable estimates of the number of required not-filers in the aggregate, it limits the reliability of analysis of filing behavior at the micro level.

Self-Employment Income

As with the Census-based method, we need to impute self-employment income to the not-filers since this income is only partially reported on third-party information returns. As stated above, the methodology assumes that not-filers look a lot like filers in terms of their propensity to have self-employment income given other characteristics, such as gender, age, region, and types and amounts of income present on third-party information returns. But, we also know from the analysis of the tax gap that self-employment income is significantly underreported on filed returns. Thus, if not-filers are generally like filers, our method underestimates the self-employment of not-filers as well. But, it could be that not-filers are less likely to have self-employment income than filers because, if they are being honest, they legitimately are not required to file. Or, given their income and other characteristics, they are more likely to have self-employment income since the relative invisibility of this form of income allows them to remain as not-filers. Whatever the conclusions about the effectiveness of the imputations in the aggregate, their effectiveness in accurately assigning self-employment income to particular taxpayers is certainly more limited. Thus, the imputation of self-employment income is another factor limiting the analysis of drivers of filing behavior at the micro level. These limitations might be partially overcome by considering the filing behavior and requirements of individuals instead of tax units and by analyzing filing drivers with and without the imputation of self-employment income.

How much of a difference do the self-employment imputations make in the VFR estimates? Figure 7 shows that when we consider only the self-employment income that is present on Forms 1099 Miscellaneous Income, the VFR is between 1.1 percent and 1.8 percent greater, with the difference being larger in recent years. Nonetheless, the trend over time in the estimates is fairly similar between the two approaches, which provides some confidence that the imputation of self-employment income mainly has the effect of lowering the level of the VFR but doesn't significantly alter the analysis of its year-to-year variation.

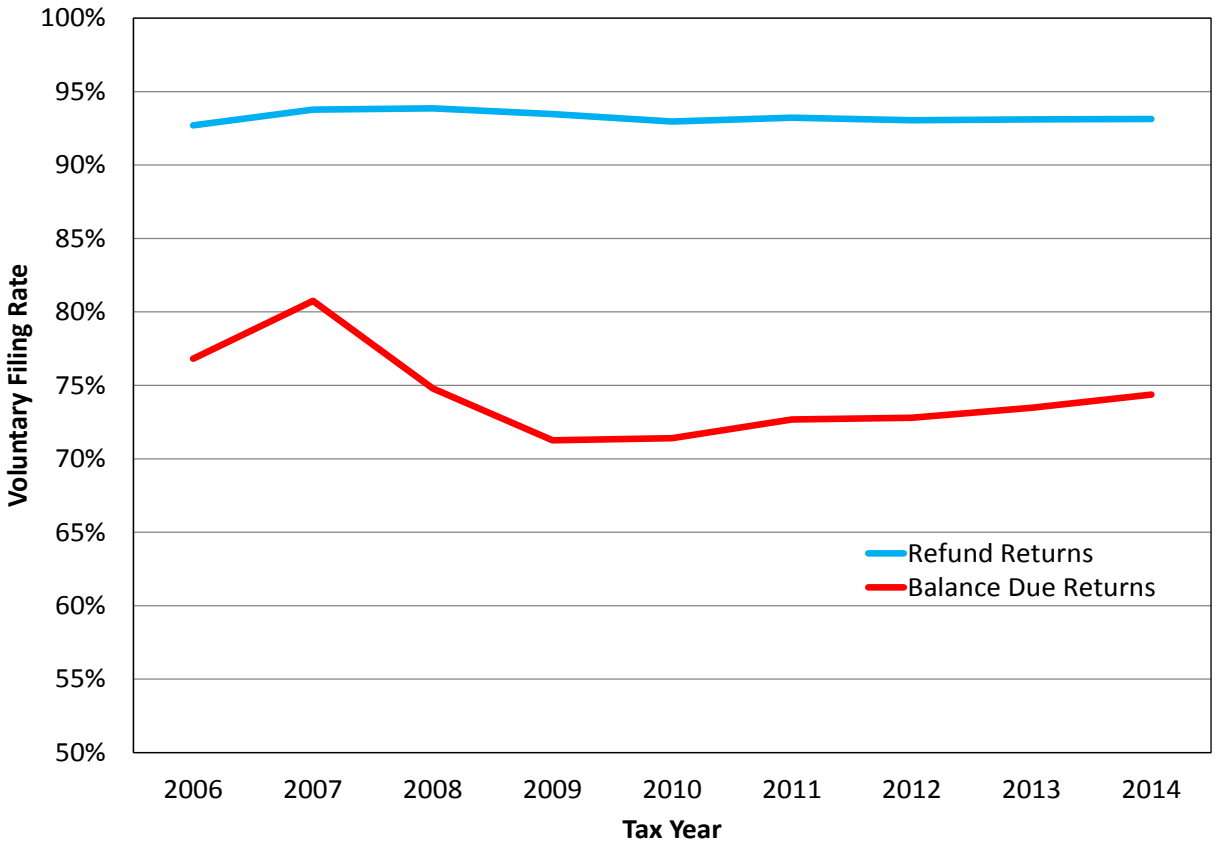
FIGURE 7. VFR With and Without Imputation of Self-Employment Income, TYs 2006–2014



Pre-Payment Position

Aside from the opportunity to obtain a tax rebate (as in the case of the 2008 Economic Stimulus Recovery Rebate Credit), another clear incentive for filing a tax return is to obtain a refund to recoup excessive tax withholding and/or overpayment of estimated taxes, or to claim refundable tax credits, such as the Earned Income Tax Credit or Additional Child Tax Credit. As Figure 8 shows, the filing rate for those owed a refund is very stable and substantially higher (about 20 percentage points greater) than the filing rate for those with a balance due. In addition, this graph shows that the effect of the economic stimulus payments in increasing the filing rate in TY 2007 was largely due to the fact that it encouraged a larger share of those with a balance due to file a tax return. Partly due to the penalties for underpayment of taxes during the year, the percent of timely filers owed a refund typically exceeds by a large magnitude those with a balance due. In TY 2014, for instance, about 78 percent of timely filers were owed a refund while only 22 percent had a balance due. As a result, the VFR is much closer to the rate of the refund group rather than the balance due group. This finding corroborates the value of the IRS' Withholding Compliance Program in ensuring both timely payment of taxes and encouraging timely filing of tax returns. Clearly, when taxpayers have sufficient taxes withheld from their paychecks to meet their tax obligations, they are also much more likely to file a tax return.

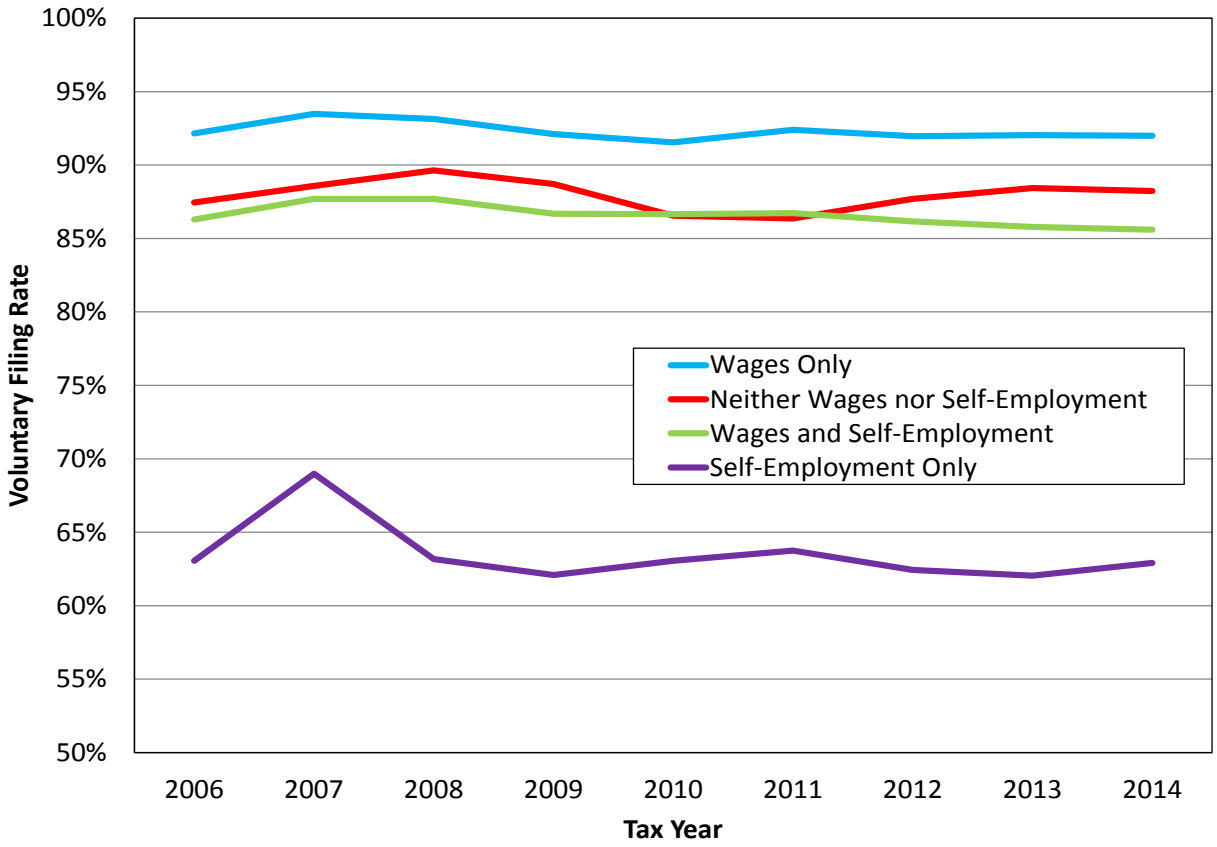
FIGURE 8. VFR by Whether Taxpayer Has a Balance Due or Owed a Refund, TYs 2006–2014



Income Visibility

Another factor that likely influences taxpayers' filing decisions is the visibility of their income to enforcement authorities. Thus, taxpayers with mostly wage income reported on Form W2 are more likely to file than those who earn most of their income from self-employment, especially when their self-employment income is not reported on third-party information returns. In addition, some taxpayers may not be aware that they are obligated to file and pay self-employment tax when their income is as low as $433. Figure 9 shows that the VFR among those whose earned income is comprised of only wage income is significantly higher (30 percentage points more) than among those with only self-employment income. The effect of the stimulus rebates in 2007 seems to have had a larger effect in encouraging filing among those with self-employment income than among those with only wage income, who were already filing at a high percentage.

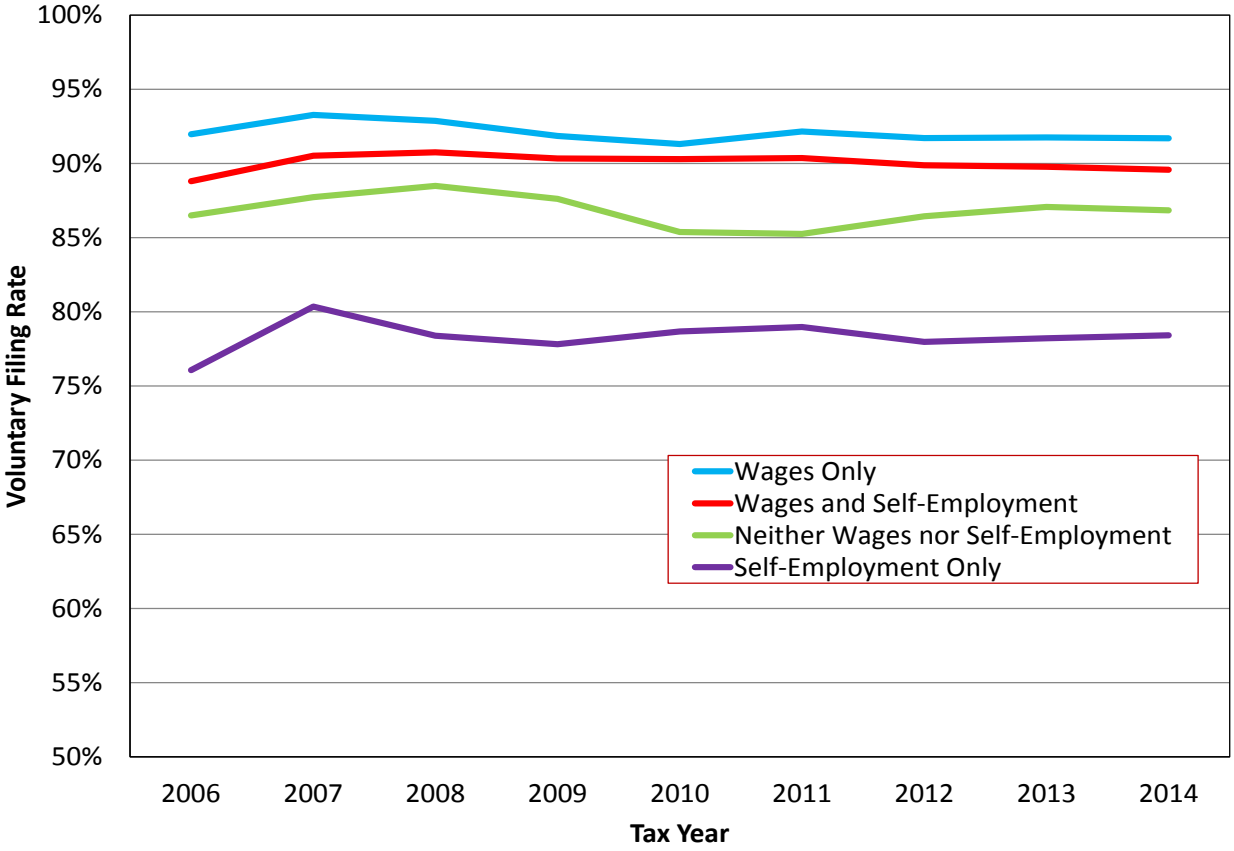
FIGURE 9. VFR by Primary Income Types (with SE Imputation), TYs 2006–2014



After imputation, more potential not-filers are assigned self-employment income. As stated above, taxpayers with self-employment income greater than $433 are required to file. However, the imputed net Schedule C amount could cause the taxpayer's gross income to exceed the relevant filing threshold even if they are below the $433 net self-employment income threshold. But, when only nonemployee compensation on Forms 1099-Misc Miscellaneous Income is considered (that is, self-employment income is not imputed), the VFR for those with just self-employment income remains significantly lower than for those with just wage income (about 78 percent instead of about 92 percent for TY 2014) (Figure 10).[17]

[17] Clearly some of those with nonemployee compensation of greater than $433 on a Form 1099-MISC do not have a filing requirement because they have sufficient expenses to reduce their net income to $433 or less. However, these same individuals, as well as other potential not-filers, likely receive self-employment income that is not reported on third-party documents. So the estimates without the self-employment imputations are likely a conservative estimate of the extent of nonfiling for this population.

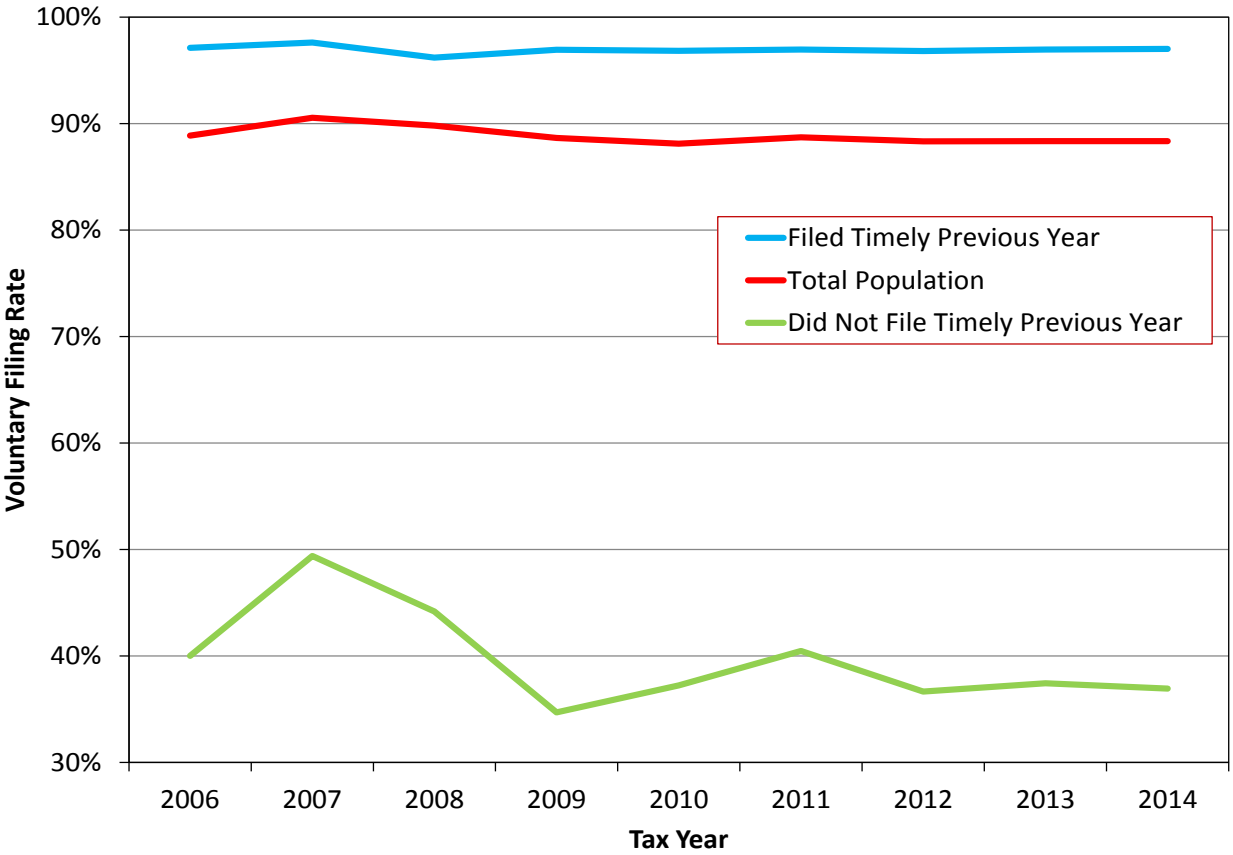
FIGURE 10. VFR by Primary Income Types (Without SE Imputation), TYs 2006–2014



Prior-Year Filing

Another presumption in the tax compliance literature is that there is significant persistence in filing behavior (Erard and Ho, 2001). On the one hand, a taxpayer who failed to file in the previous tax year has less incentive to file in the current year because he likely perceives that doing so increases the probability that his past filing noncompliance will be discovered. On the other hand, a taxpayer who did file a return in the previous year risks raising the suspicion of the tax authority if he does not file in the current year. In addition, once a taxpayer starts filing a tax return, it can become a matter of habit. Thus, it is not very surprising that, in fact, the VFR is much higher among those who filed timely in the previous tax year. For those who filed timely the prior year, the VFR hovers between 96 percent and 97 percent, while for those who did not file timely in the prior year, the VFR ranges between 35 percent and 50 percent (Figure 11).

FIGURE 11. VFR for Taxpayers Who Filed Timely in the Previous Year vs. Those Who Did Not, TYs 2006–2014

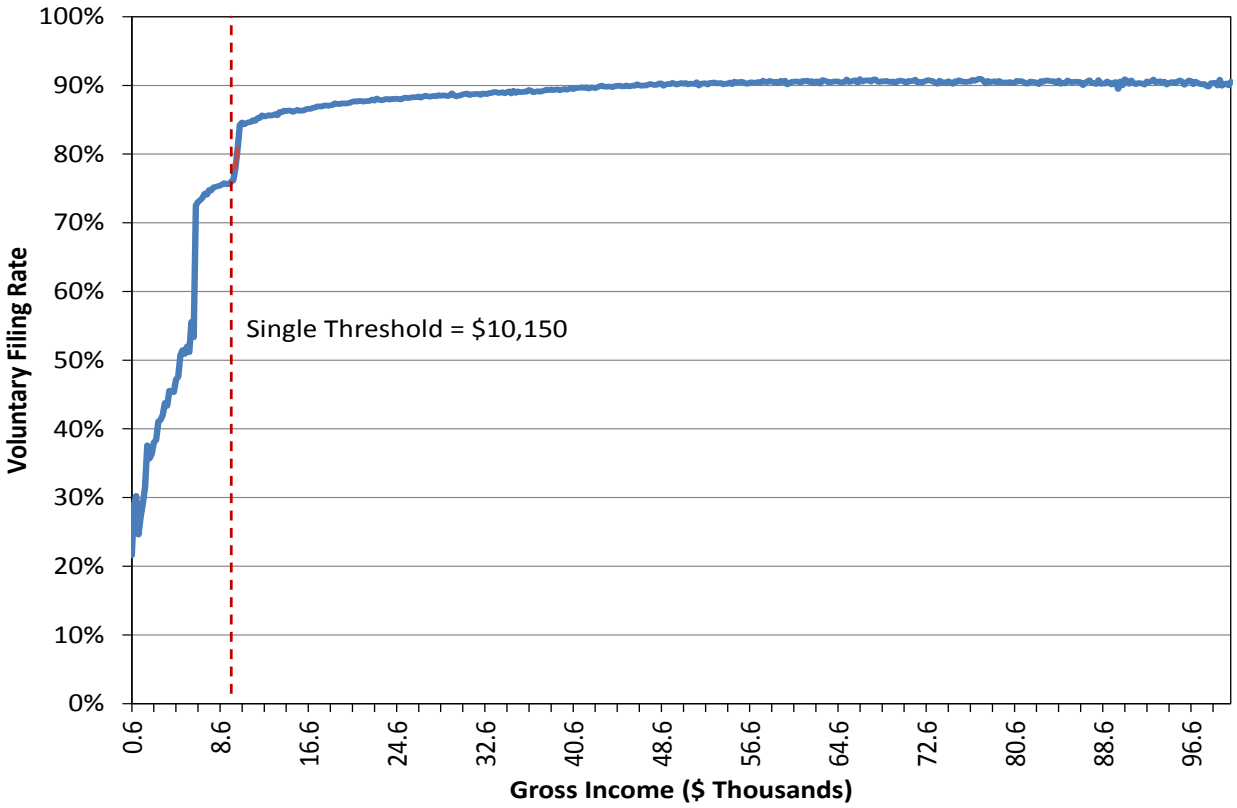


Amount of Income

Another factor that could be expected to make a taxpayer more likely to file is the level of income relative to the pertinent filing threshold. Taxpayers with gross income below the filing threshold, but with enough Schedule C income that they owe self-employment tax, might not be aware that they are obligated to file a tax return. When gross income barely exceeds the threshold, taxpayers may believe that they are below tolerance for IRS enforcement, especially if they would be owed a refund if they filed a return. A taxpayer whose income is only slightly above the filing threshold may owe very little tax or be due a refund that is too small to offset the financial costs and burden of filing a return.

Figures 12 and 13 tend to support these expectations about how filing relates to income. First, the VFR is very low for those well below the gross income filing threshold, but required to file because of self-employment tax or because they are claimed as dependents on another return. Second, the VFR jumps sharply as gross income nears the threshold and then gradually increases logarithmically as gross income increases beyond the threshold. This pattern holds true for both single and married taxpayers, but the income effect seems more prolonged and stronger in the case of married taxpayers.

FIGURE 12. VFR of Single Taxpayers (Under 65 Years Old) by Gross Income, TY 2014



Age

Filing rates may also vary by age, but the nature of this influence is not clear a priori. Changes in cultural norms might affect age groups differentially and the composition and level of income, filing status, and numbers of dependents may change as taxpayers age, which could affect taxpayers' propensity to file. Given the interrelationship of age with many filer characteristics,teasing out its effects would likely require a multivariate analysis. But, the plots in Figures 14, 15 and 16 show the relationship between the VFR and age for all taxpayers as well as for single and married taxpayers separately, which at least raise questions for further research. Figure 14 suggests that the VFR is relatively high for young taxpayers, declines as taxpayers reach middle age, increases again as taxpayers near retirement, and then falls off in later years. This pattern is accentuated for single taxpayers (Figure 15) and modulated for married taxpayers (Figure 16). Without controlling for other factors, married taxpayers generally have a higher filing rate than single taxpayers. It would be interesting to explore the factors other than age that are driving the apparent decline in the voluntary filing rate in the middle age ranges and leading to lower apparent filing rates among elderly taxpayers.

One factor that we thought might contribute to lower apparent filing rates for the elderly is the fact that amounts reported on 1099Rs as taxable pensions are not necessarily subject to taxation.[18] This could have resulted in an overestimation of the numbers of nonfiling elderly taxpayers with a filing requirement and an underestimate of the VFR. However, using the random audit data from the National Research Program to correct for the misidentification of pension income as taxable on 1099Rs, we found that this problem was a not a major factor in the apparent VFR decline for these taxpayers.[19]

[18] About 18 percent of taxpayers with a taxable pension amount on a 1099R were found in NRP audits to have no taxable pension income. For those who did have verified taxable pension amounts on their Form 1040s, the amounts on average were about 10 percent lower than the amounts reported on the 1099Rs.

[19] Other than measurement limitations, several other factors, including a greater prevalence of health problems, lower awareness of filing rules, and lower computer literacy could be driving the lower filing rates among the elderly.

FIGURE 13. VFR of Married Taxpayers (Under 65 Years Old) by Gross Income, TY 2014

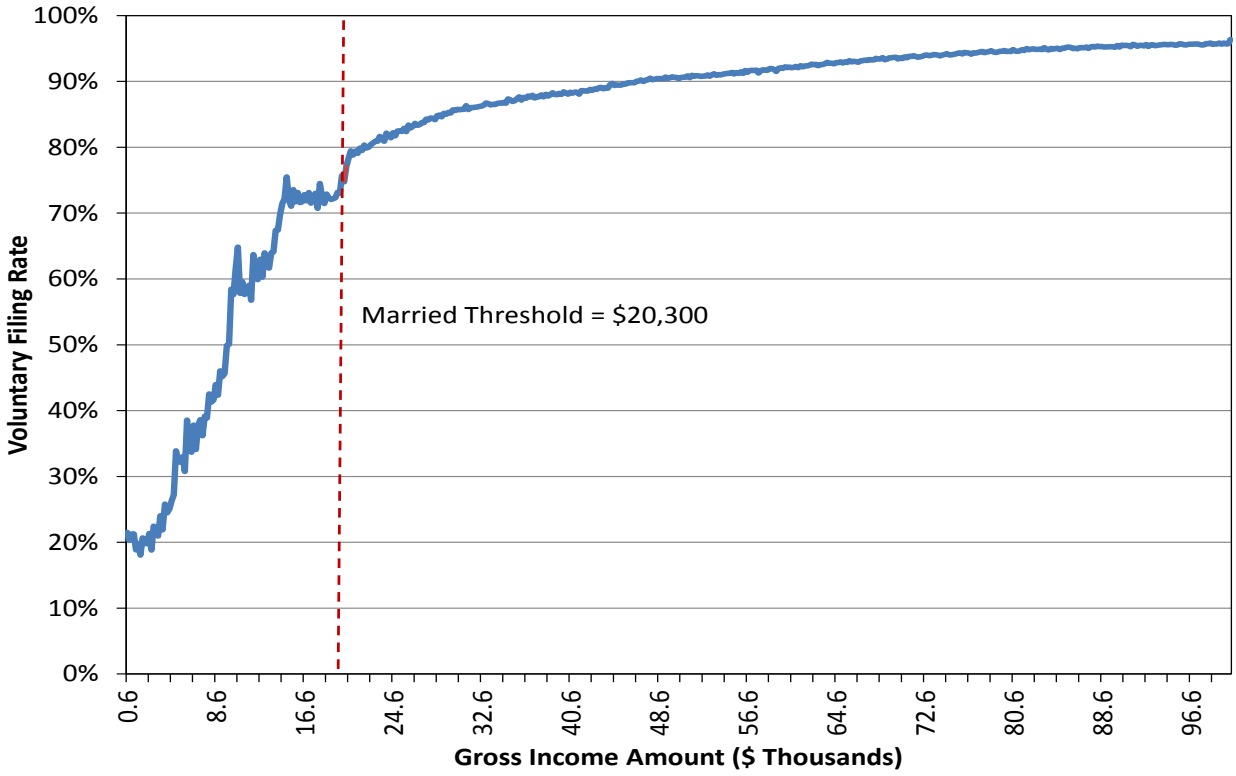


FIGURE 14. VFR by Age, TY 2014

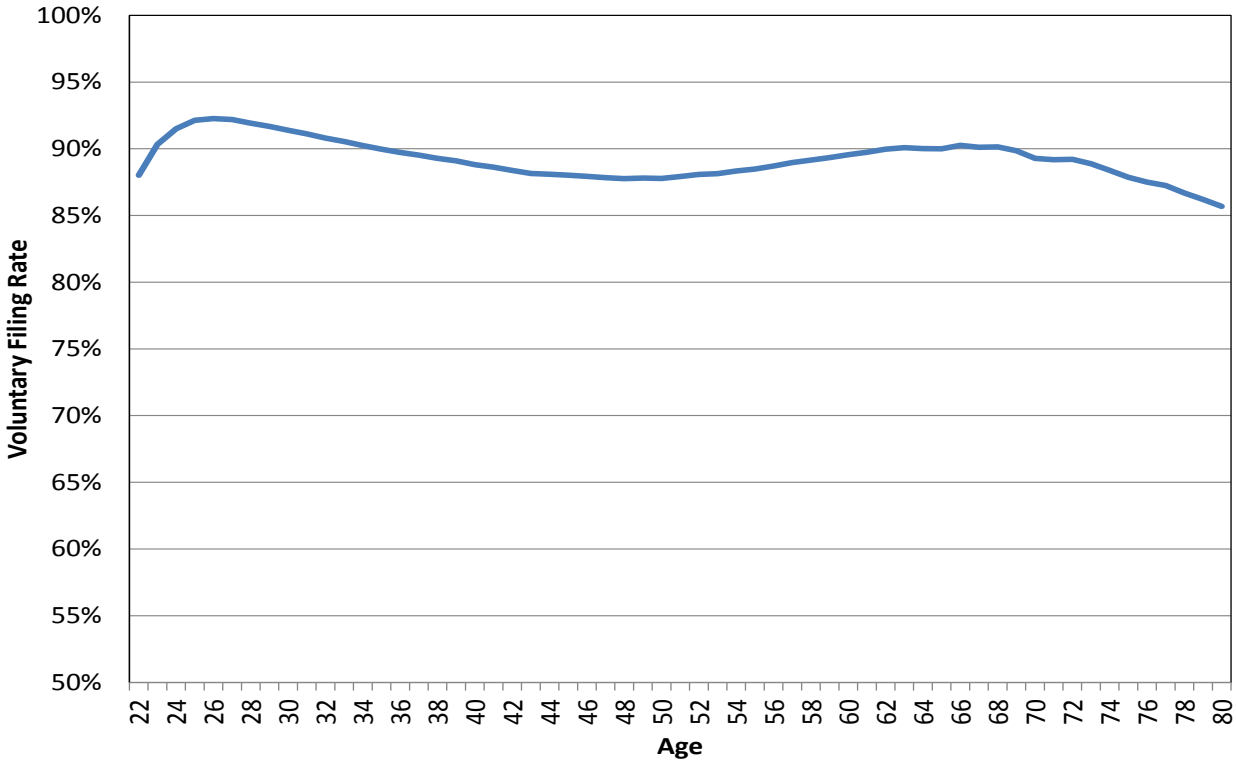


FIGURE 15. VFR by Age, Single Taxpayers, TY 2014

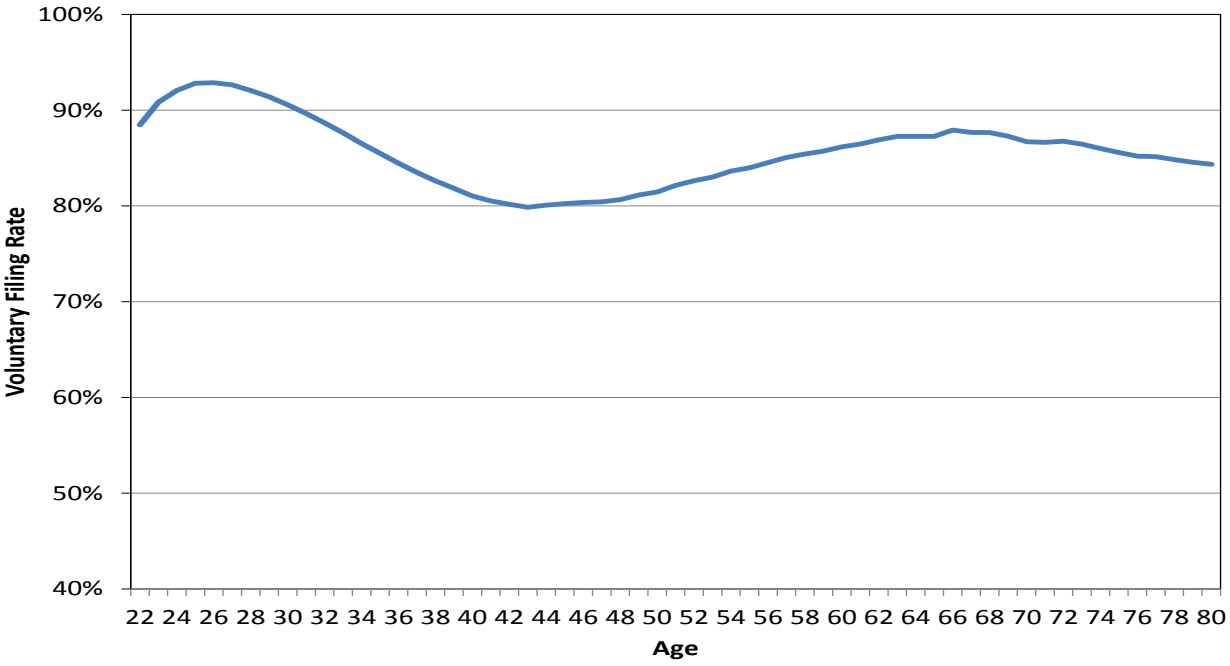
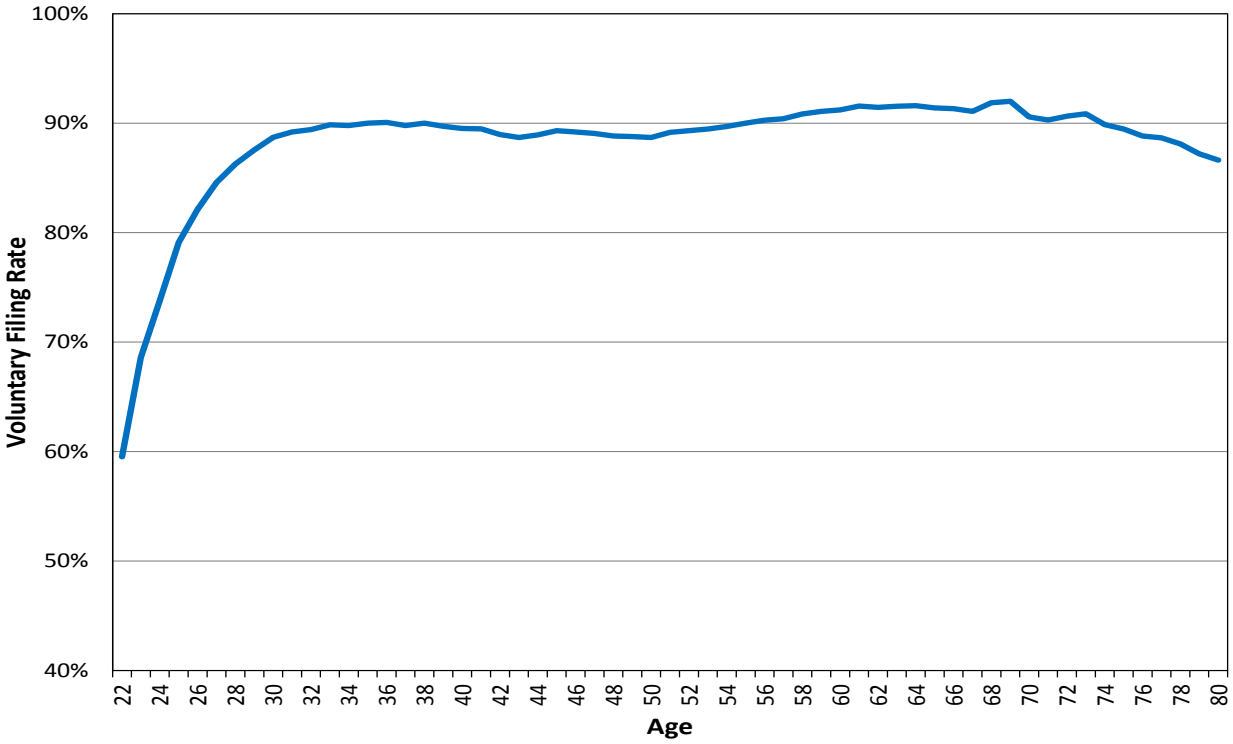


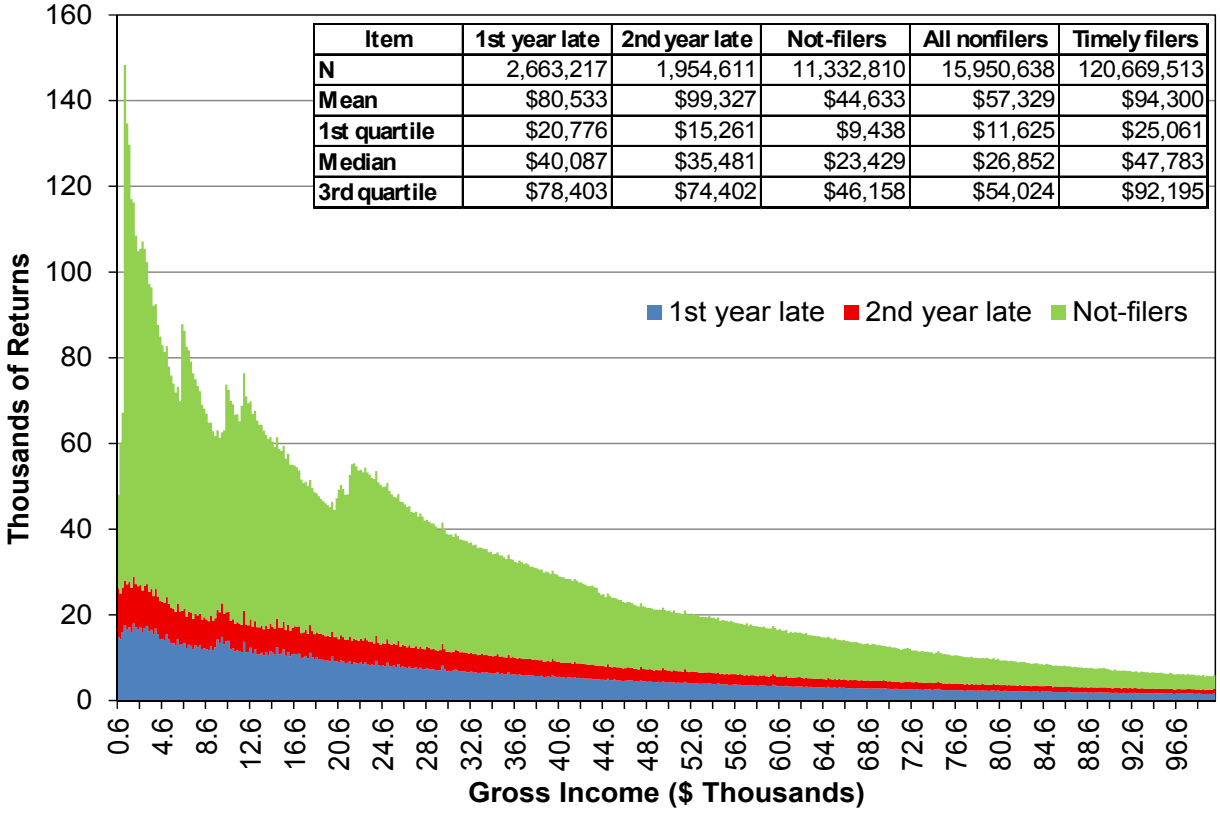
FIGURE 16. VFR by Age, Married Taxpayers, TY 2014



Income and Tax Comparisons With Timely Filers

The next several figures examine some of the distributional characteristics of nonfilers compared with timely filers. Given the fact that the filing rate is lower for taxpayers close to the filing threshold, it is not surprising that the average gross income of required nonfilers is lower than that of required timely filers (Figures 17 and 18). For nonfilers, mean gross income in TY 2014 was $57,329, while for timely filers it was $94,300. The mean gross income for late filers is very close to that of timely filers, while that of not-filers is much smaller ($44,633). The returns that come in during the first year after the tax year tend to report less income and to be less unequally distributed than returns that are filed in the second year after the tax year. The distribution of gross income for not-filers is skewed just as much as for timely filers, with the upper one-quarter of the distribution responsible for about one-half of total gross income.

FIGURE 17. Distribution of Gross Income, Late and Not-Filers, TY 2014



Because not-filers tend to make less income than timely and late filers, they are also estimated to owe less tax. While timely filers report an average tax liability of $11,530, not-filers are estimated to owe an average of $7,957. The distribution of total tax is considerably more skewed for not-filers than it is for timely filers (Figure 19).

FIGURE 18. Distribution of Gross Income, Timely Filers (TY 2014)

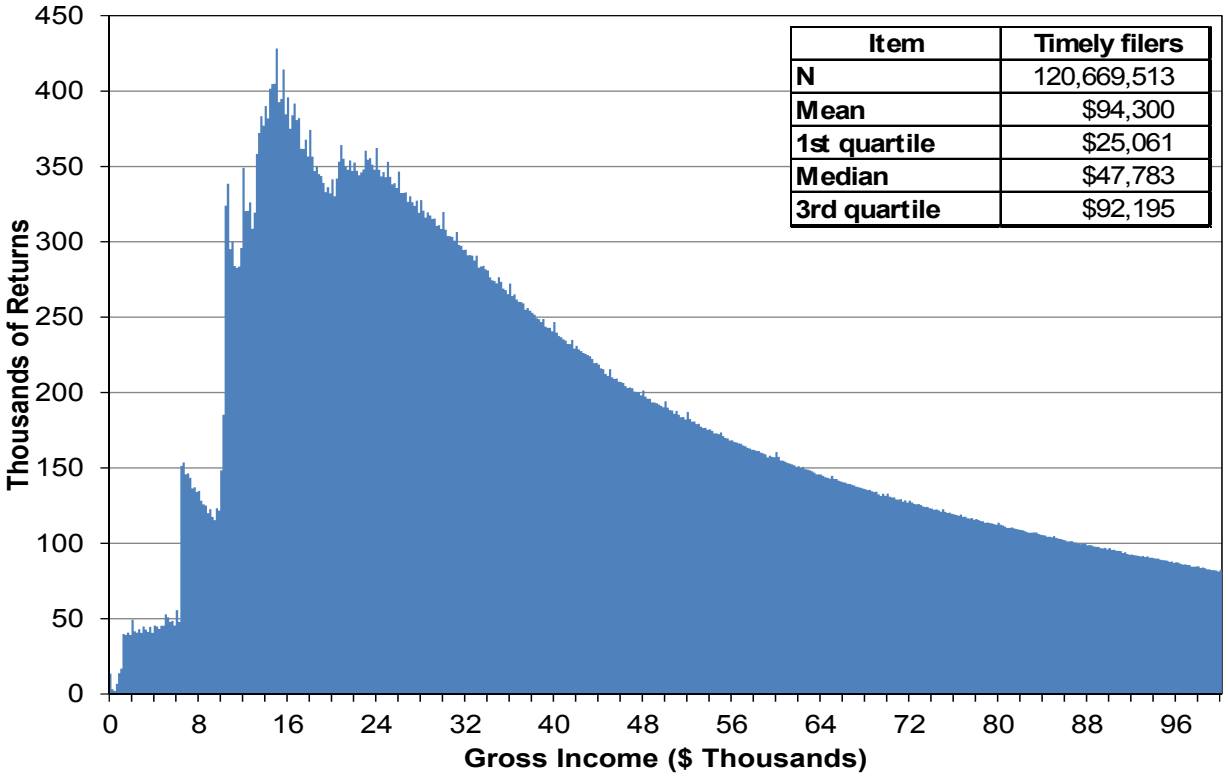
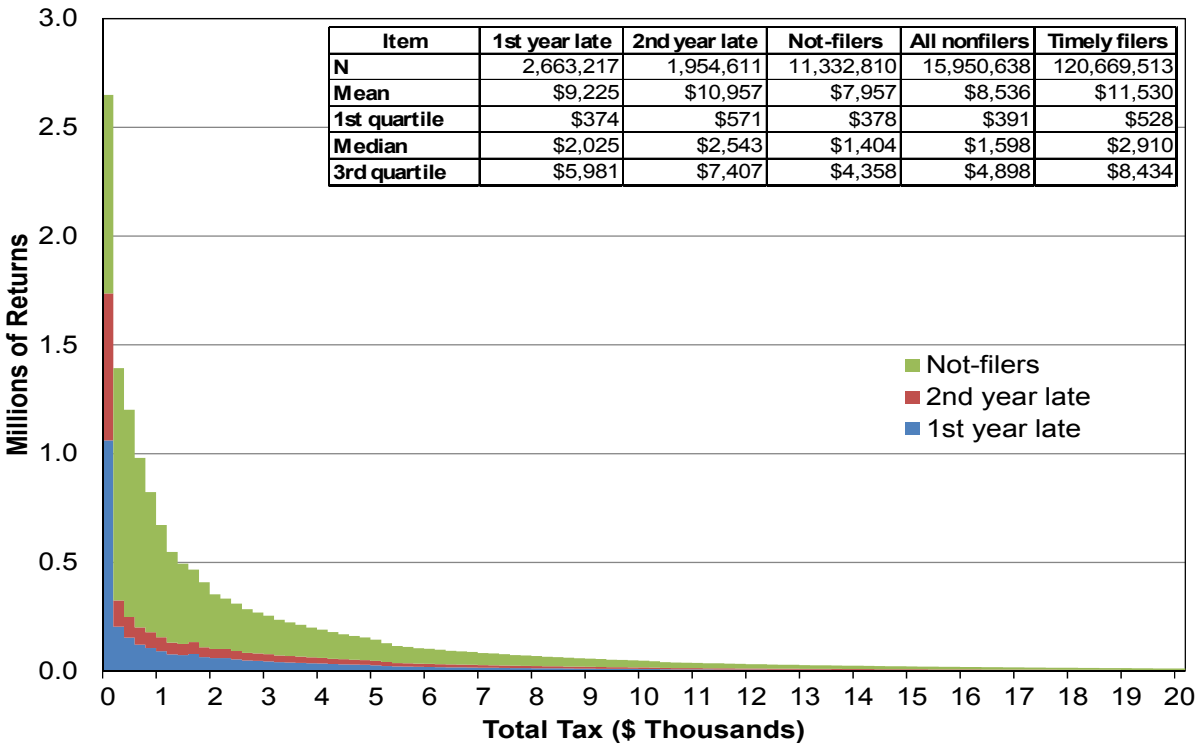
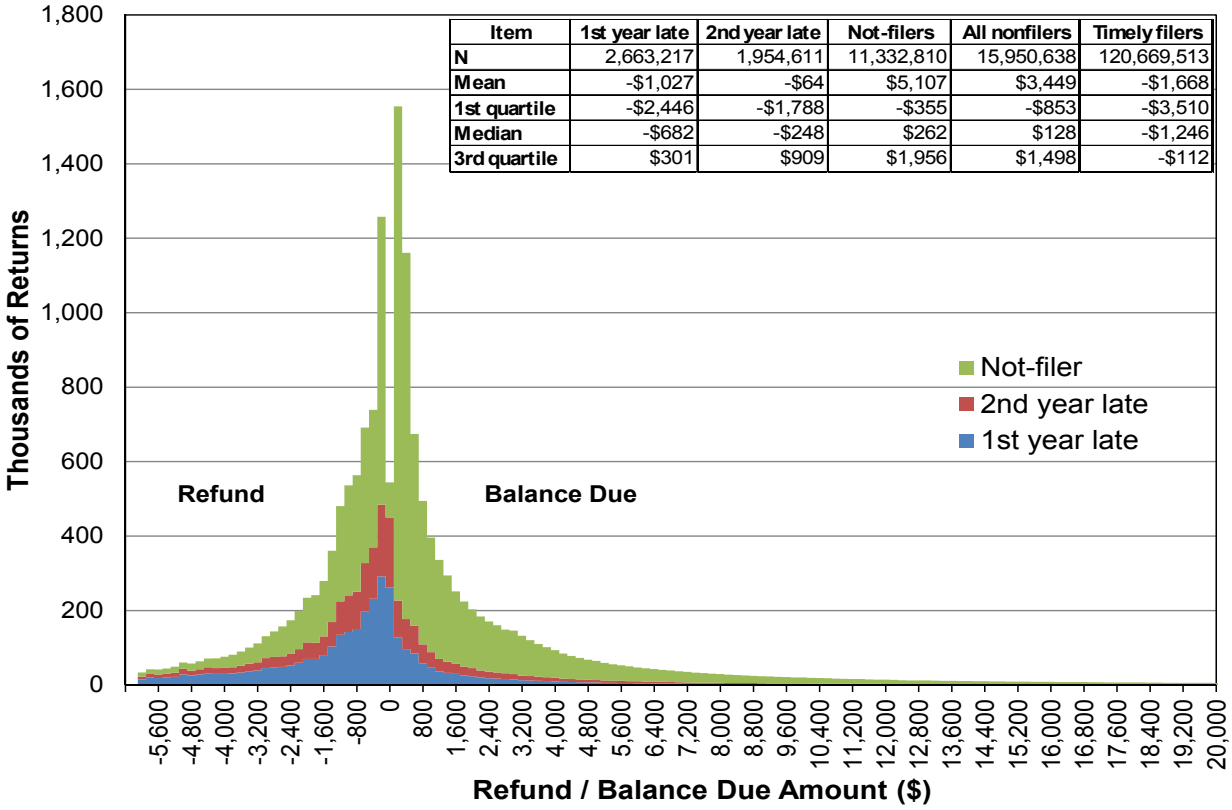


FIGURE 19. Distribution of Total Tax, Late Filers and Not-Filers, TY 2014



Given that about 78 percent of taxpayers filing a timely return are owed a refund, it is not surprising that on average the tax returns of timely filers have a negative balance due (Figure 20). This is also true for late filers, although in the case of second year late filers, refund and balance due returns are closer in number than for returns filed late in the first year following the tax year. However, not-filers are estimated on average to owe over $5,000. As discussed above, the presence of a tax debt is likely a key reason why such taxpayers choose not to file a return.

FIGURE 20. Distribution of Refund/Balance Due Amount, 1st Year Late, 2nd Year Late, and Not-Filers, TY 2014



Benefits of the VFR Research

Our efforts to develop a VFR measure using the Census-IRS matched data and by using just IRS data have resulted in important benefits. In the attempt to use the Census-IRS matched data, we found that even after imputing additional income to the Census records following the previous methodology, the estimated number of required taxpayers was still dramatically lower than the estimate from IRS data. This indicates that there are likely other forms of income that are insufficiently reported in the CPS survey that contribute to this undercount. Thus, the work on the VFR has revealed additional gaps in the CPS income data relative to the income reported in IRS data. This led to some efforts using the matched IRS-CPS data to try to fill these gaps, but more work along these lines is needed to ensure that the resulting distribution of income in the CPS is comparable to that found in the IRS population data.

The work on the IRS administrative VFR has resulted in a more accurate VFR that allows us to more carefully examine the drivers of filing behavior and to explain the fluctuations in the VFR over time. There are benefits to being able to explore the role of the numerator and denominator together—rather than just the numerator—in affecting fluctuations in the VFR. The time series data resulting from this work could, with some limitations, potentially be used in a multivariate analysis of the drivers of filing behavior.

Future Work

For the IRS-based VFR presented in this paper, we assume that the income amounts and filing status as reported and processed on the return are correct. However, we know that it is likely that a significant number of single filers incorrectly claim Head of Household status on their tax returns, and by so doing, many of them appear to not have a filing requirement, when in fact they do. Thus, one potential modification of the measure would be to use the audit data from the IRS National Research Program to impute corrected filing statuses to those who claim Head of Household status. This would tend to increase the number of required returns filed on time, and therefore increase the VFR. Another potential refinement of the IRS-based VFR would be to add to the filed returns the income amounts reported on third-party information documents but not on the returns. This change might also add to the count of required returns that were filed timely, while likely adding proportionately fewer to the category of late returns.

Greater leverage in understanding the drivers of filing might be obtained by using the IRS administrative data approach in a longitudinal probit analysis of the decision to file. Given the inaccuracies introduced at the micro-level in the imputation of tax units for the not-filers, one approach might be to analyze filing at the individual level, rather than the tax unit level. Another approach would be to improve the imputation of tax units by drawing on information from prior year returns and other sources, such as Social Security Administration data. Since the self-employment imputations also introduce inaccuracies at the micro-level, the size and significance of the parameter estimates could be compared with and without both sets of imputations.

It would be valuable to explore the factors driving variation in filing rates across different demographic groups. A better understanding of these drivers might suggest outreach programs or targeted enforcement efforts that could encourage filing for groups having lower filing rates.

Another area of future work is to further explore the potential of the expanded IRS-Census matched dataset in supporting the development of a reliable alternative VFR measure, as well as to further examine the drivers of filing behavior. With complete IRS data in the Census environment, it should be possible to explore more directly the deficiencies in the matching process as well as in the reporting of income on CPS records. The matched data also provides an opportunity to undertake an analysis of the drivers of filing without the same limitations of knowledge about the structure of the nonfiling tax units.

References

Erard, Brian and Chih-Chin Ho (2001). "Searching for Ghosts: Who Are the Nonfilers and How Much Tax Do They Owe?," Journal of Public Economics. Vol. 81 pp. 25–50.

Erard, Brian, Pat Langetieg, Mark Payne, and Alan Plumley (2014), "Missing Returns vs. Missing Income: Estimating the Extent of Individual Income Tax Filing Noncompliance From IRS and Census Data," Proceedings of the National Tax Association Annual Conference, November 2014.

Erard, Brian, Mark Payne, and Alan Plumley (2012), "Advances in Nonfiling Measures," 2012 IRS Research Bulletin, Internal Revenue Service, Publication 1500, Washington, DC, pp. 79–89.

Langetieg, Pat, Mark Payne, and Alan Plumley (2016), "The Individual Income Tax and Self-Employment Tax Nonfiling Gaps for Tax Years 2008–2010," 2016 IRS Research Bulletin, Internal Revenue Service, Publication 1500, Washington, DC, pp. 38–59.

Plumley, Alan (1996), "The Determinants of Individual Income Tax Compliance: Estimating the Impact of Tax Policy, Enforcement and IRS Responsiveness," Internal Revenue Service, Publication 1916 (Rev. 11-96), Washington, DC, 1996.

5

∇

Appendix

Conference Program

7th Annual IRS-TPC Joint Research Conference on Tax Administration
Urban Institute, 2100 M Street, NW, Washington, DC
June 21, 2017

Program

8:30 – 9:00 Check in

9:00 – 9:15 Opening

Welcome ***Eric Toder*** *(Co-Director, Tax Policy Center) and*
Ben Herndon *(Director, IRS Office of Research, Applied Analytics, and Statistics (RAAS))*
John A. Koskinen (IRS Commissioner)

9:15 – 10:45 Identifying Corporation Tax Avoidance

Moderator: *Eric Toder (Tax Policy Center*)

- Using IRS Data to Identify Income Shifting Firms, *Lisa De Simone (Stanford University), Lillian F. Mills (The University of Texas-Austin), and Bridget Stomberg (The University of Georgia)*
- Income Shifting by U.S. Multinational Corporations. *Ted Black (IRS, RAAS), Amy Dunbar (University of Connecticut), Andrew Duxbury (James Madison University), and Thomas Schultz (Western Michigan University)*
- The Economic Effects of Special Purpose Entities on Corporate Tax Avoidance, *Paul Demeré*, Michael P. Donohoe, and Petro Lisowsky (University of Illinois at Urbana-Champaign)

Discussants: *Tim Dowd (Congressional Joint Committee on Taxation Staff) and Eric Toder (Tax Policy Center)*

10:45 – 11:00 Break

11:00 – 12:30 Realizing the Potential of Tax Enforcement

Moderator: *Ron Hodge (IRS, RAAS)*

- How Do IRS Resources Affect the Tax Enforcement Process?, Ron Hodge (IRS, RAAS), *Michelle Nessa (Michigan State University), Casey Schwab (Indiana University), and Bridget Stomberg and Erin Towery (University of Georgia)*
- Tax Audits and Tax Compliance—Evidence from Italy, *Elena D'Agosto, Marco Manzo, Alessandro Modica and Stefano Pisani (Ministero dell'Economia e delle Finanze, Government of Italy)*
- Valuing Unpaid Tax Assessments—Estimating Long-Run Collectability Using an Econometric Approacht, *Dan Howar and Maryamm Muzikir (IRS, SB/SE), Alex Turk (IRS, RAAS), and Eric Henry (IRS, CFO)*

Discussants: *Michael Udell (District Economics Group)*

1:00 – 1:30 Keynote Speaker

Peter Merrill (National Economics & Statistics, PricewaterhouseCoopers LLP)

1:30 – 3:00 The Role of Incentives in Individual Compliance

Moderator: *Saima Mehmood (IRS, Wage & Investment Division Research)*

- Impact of Filing Reminder Outreach on Voluntary Filing Compliance for Taxpayers with a Prior Filing Delinquency, *Stacy Orlett, Maryamm Muzikir, and Vicki Koranda (IRS, SB/SE), and Rizwan Javaid and Alex Turk (IRS, RAAS)*
- Charitable Contributions of Conservation Easements, *Adam Looney (The Brookings Institution)*
- Tax Preparers, Refund Anticipation Products, and EITC Compliance, *Maggie R. Jones (U.S. Census Bureau)*

Discussant: *Janet Holtzblatt (Congressional Budget Office)*

3:00 – 3:10 Break

3:10 – 4:40 Creative Use of Non-Tax Data Sources

Moderator: *Emily Lin (U.S. Treasury Office of Tax Analysis)*

- Supplementing IRS Data with External Credit Report Data in Employment Tax Predictive Models, *Curt Hopkins and Ken Su (IRS, SB/SE)*
- Better Identification of Potential Employment Tax Noncompliance Using Credit Bureau Data, *Saurabh Datta, Patrick Langetieg, and Brenda Schafer (IRS, RAAS)*
- Estimating the Effects of Tax Reform on Compliance Burdens, *Daniel Berger and Eric Toder (Tax Policy Center), and Victoria Bryant, John Guyton, and Patrick Langetieg (IRS, RAAS)*
- Counting Elusive Nonfilers Using IRS Rather Than Census Data, *Patrick Langetieg, Mark Payne, and Alan Plumley (IRS, RAAS)*

Discussants: *Adam Isen and Emily Lin (U.S. Treasury Office of Tax Analysis)*

4:40 – 4:50 Wrap-up

Ben Herndon (Director, IRS Office of Research, Applied Analytics, and Statistics)