Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

11th Annual IRS/TPC Joint Research Conference on Tax Administration

June 24, 2021

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

11th Annual IRS/TPC Joint Research Conference on Tax Administration

Welcome

Eric Toder	Institute Fellow, Urban Institute, and Codirector, Urban-Brookings Tax Policy Center
Barry Johnson	Acting Chief Research and Analytics Officer, IRS

11th Annual IRS/TPC Joint Research Conference on Tax Administration

Opening Remarks

Chuck Rettig — Commissioner of Internal Revenue

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

Session 1. Improving Individual Taxpayer Compliance

Moderator:	***Robert McClelland*** *Tax Policy Center*
Audit Contagion? Investigating the General Indirect Effect of Audits Through Tax Preparer Networks	***Kyle Furlong*** *MITRE Corporation*
Do Collateral Sanctions Work? Evidence from the IRS' Passport Certification and Revocation Process	***Paul Organ*** *University of Michigan*
EITC Noncompliance: Examining the Roles of the Dynamics of EITC Claims and Paid Preparer Use	***Alexander Yuskavage*** *Treasury Office of Tax Analysis*
Discussant:	***Tatiana Homonoff*** *New York University*

Audit Contagion?: Investigating the General Indirect Effect of Audits through Tax Preparer Networks

Kyle Furlong[1], Ellen Badgley[1], Lucia Lykke[1], Leigh Nicholl[1], and Alan Plumley[2]

June 24, 2021

[1]The MITRE Corporation
[2]Internal Revenue Service

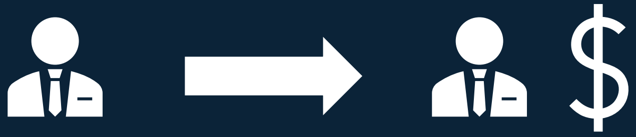
Internal Revenue Service

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD™

What is the indirect effect of tax enforcement?

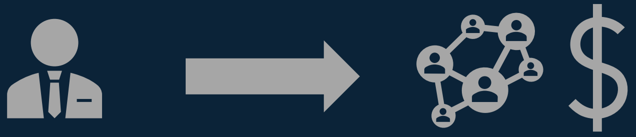
Specific Indirect Effect

Experiencing an enforcement activity changes subsequent year behavior for that same taxpayer



General Indirect Effect

Experiencing an enforcement activity changes subsequent year behavior for other taxpayers around the audited taxpayer

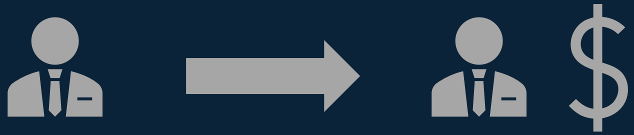


MITRE

What is the indirect effect of tax enforcement?

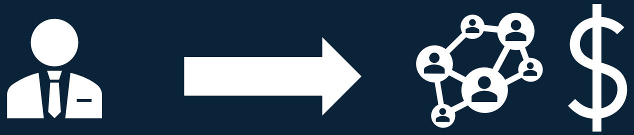
Specific Indirect Effect

Experiencing an enforcement activity changes subsequent year behavior for that same taxpayer



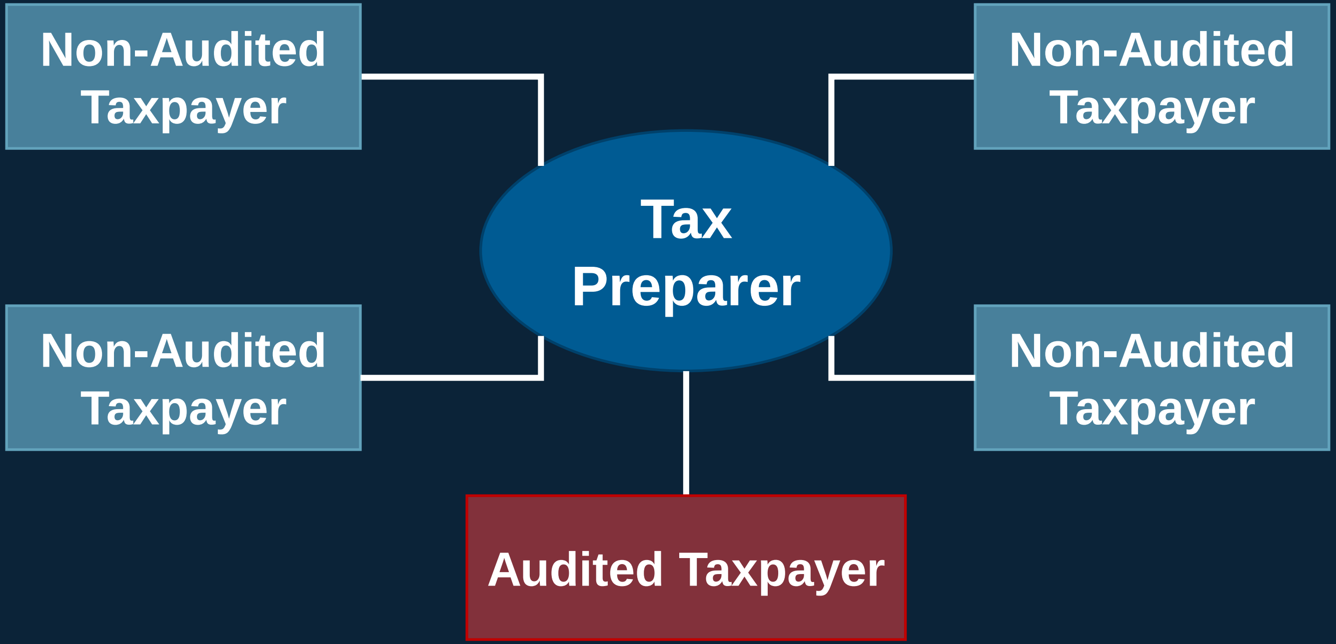
General Indirect Effect

Experiencing an enforcement activity changes subsequent year behavior for other taxpayers around the audited taxpayer



MITRE

Can we demonstrate from operational audit data that tax preparer networks facilitate a general indirect effect?



Background and Motivation

- The general indirect effect has recently been observed between **firms** that share a tax preparer when one firm is visited by a revenue officer (Boning *et al*. 2020).
- Tax preparers can exert influence over **individual** taxpayers' reporting (Batta *et al.* 2019; Erard 1993; Klepper *et al.* 1991).
- **But we don't know…**
  - If the general indirect effect of an audit propagates between **individual** taxpayers who share a tax preparer.
  - If taxpayer-tax preparer relationships change over time, which could affect our ability to estimate this indirect effect.

Research Questions

1. To what extent do taxpayers remain with their tax preparers over time?

- We investigate the dynamics of the taxpayer-tax preparer relationship over 7 years.
- We assume that the indirect effect arises because tax preparers interact with other taxpayers when an audit begins.

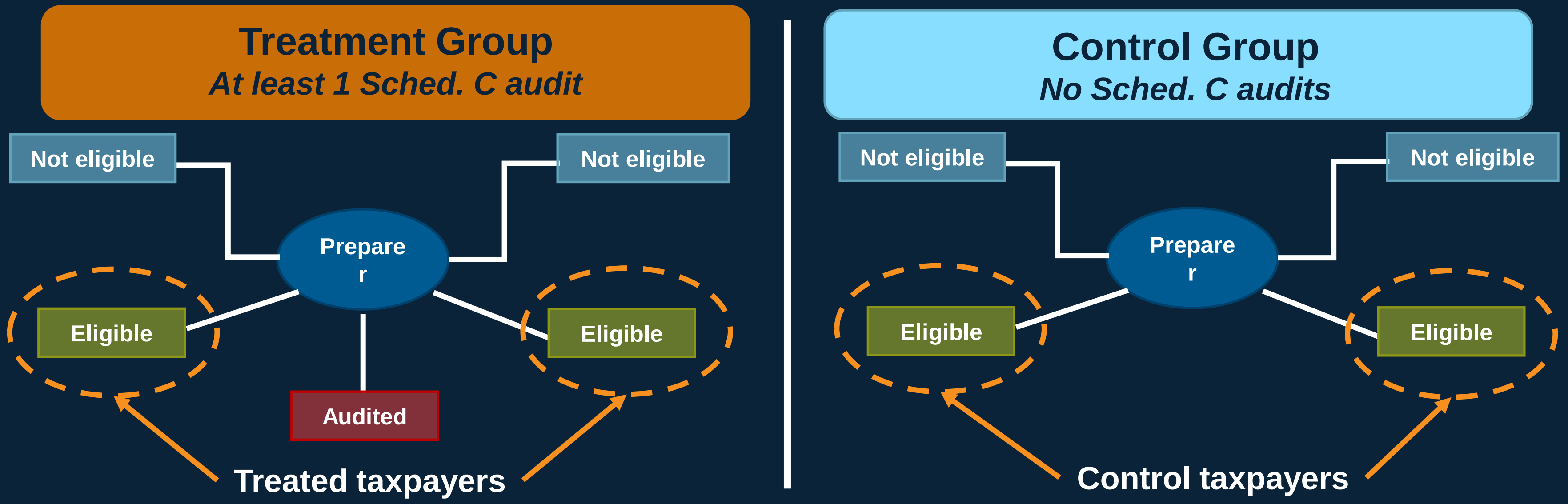
2. Do taxpayers in audited preparer networks differ in their total tax reporting over time compared to taxpayers in non-audited preparer networks?

- We compare tax reporting among *audit-eligible taxpayers* in preparer networks *who are not themselves audited*

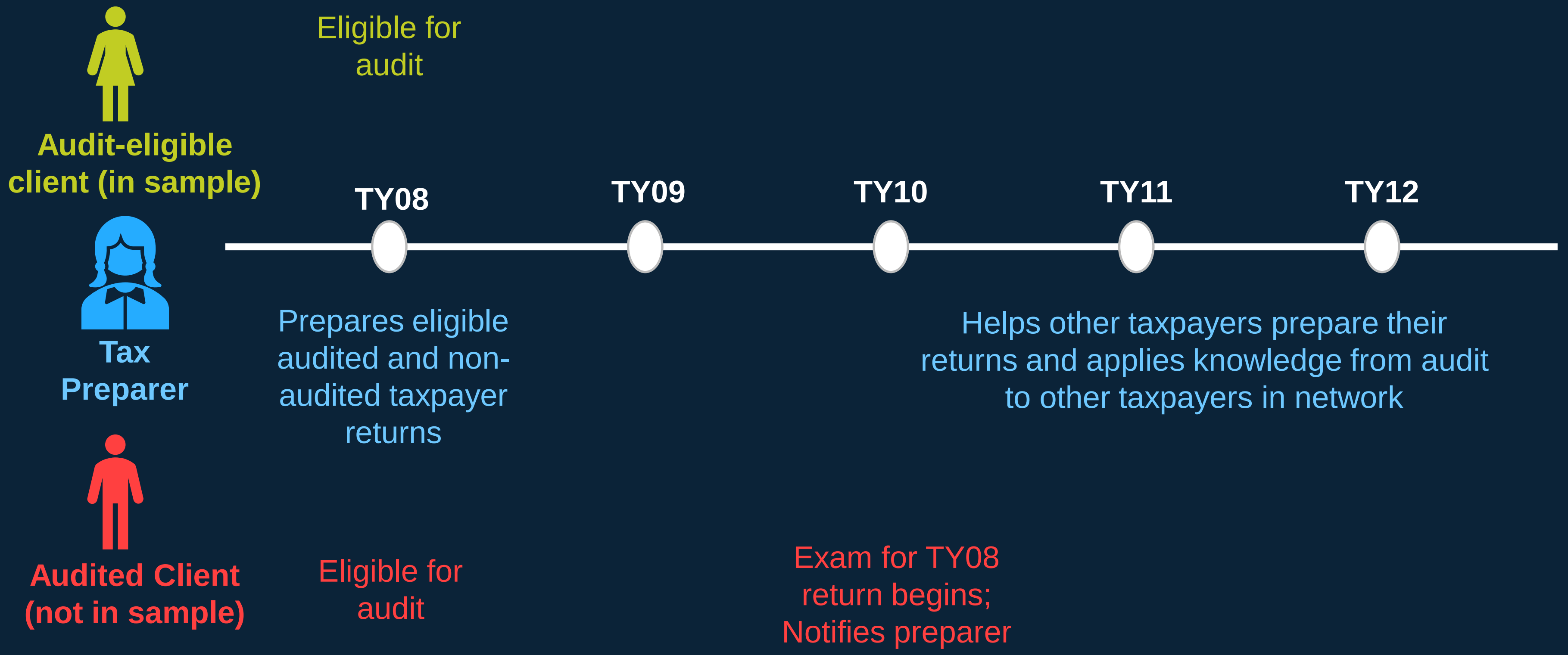
MITRE

Analytical Sample

- Tax preparers with/without **Schedule C Correspondence audits** in their network
- *Baseline year* = Tax Year (TY) the taxpayer entered the sample due to audit eligibility
- Treatment group: **unaudited taxpayers** in the **audited tax preparer network**
- Control group: **unaudited taxpayers** in the **unaudited tax preparer network**



Hypothetical Timeline of a "Treated" Network



MITRE

Hypothetical Timeline of a “Treated” Network

Audit-eligible client (in sample)

Eligible for audit

Changes reporting behavior

Tax Preparer

TY08

TY09

TY10

TY11

TY12

Prepares eligible audited and non-audited taxpayer returns

Exams take on average 2-3 years to open after filing

Helps other taxpayers prepare their returns and applies knowledge from audit to other taxpayers in network

Audited Client (not in sample)

Eligible for audit

Exam for TY08 return begins; Notifies preparer

MITRE

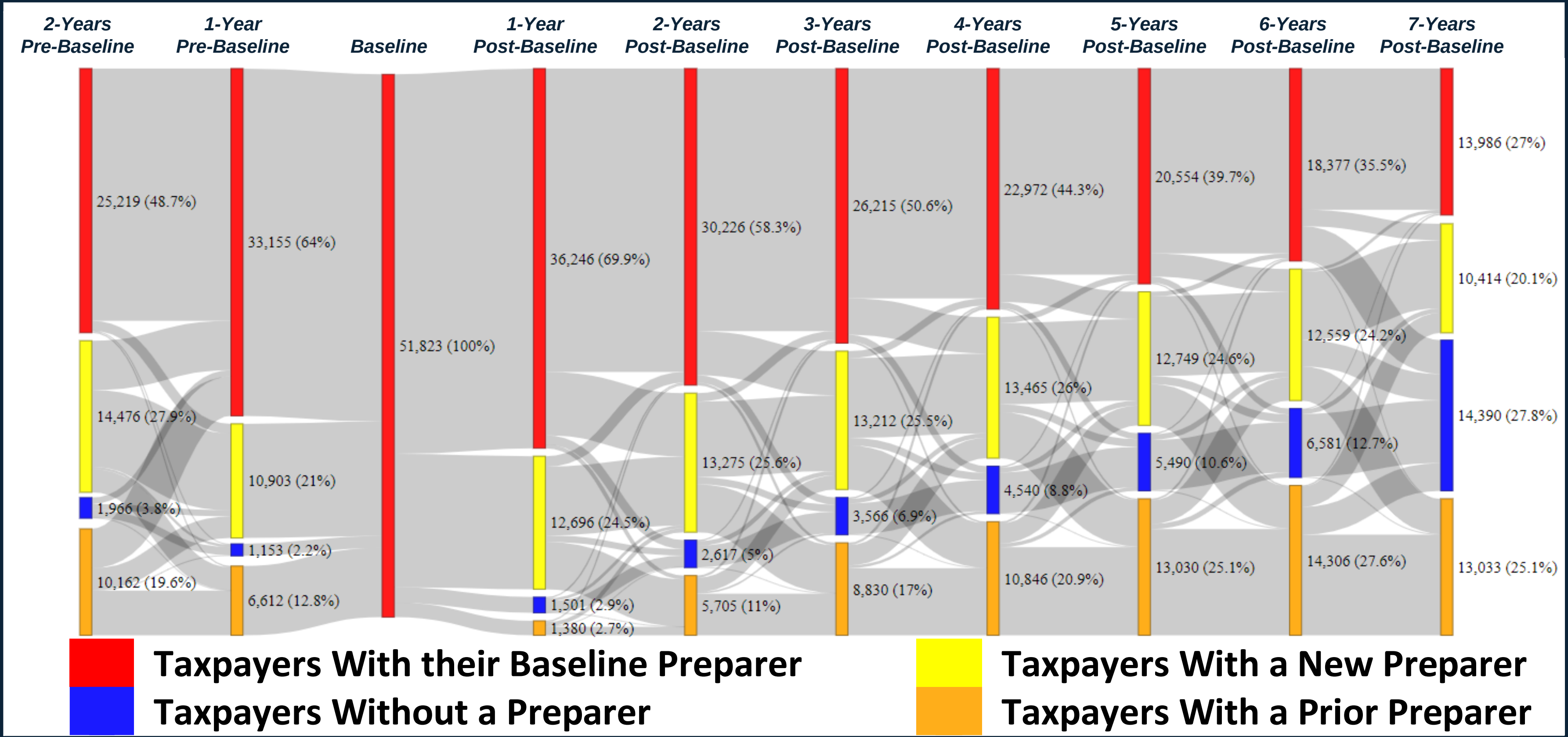
Taxpayer-Tax Preparer Dynamics: Data Visualization

Taxpayer-Tax Preparer Dynamics with Sankey Diagrams

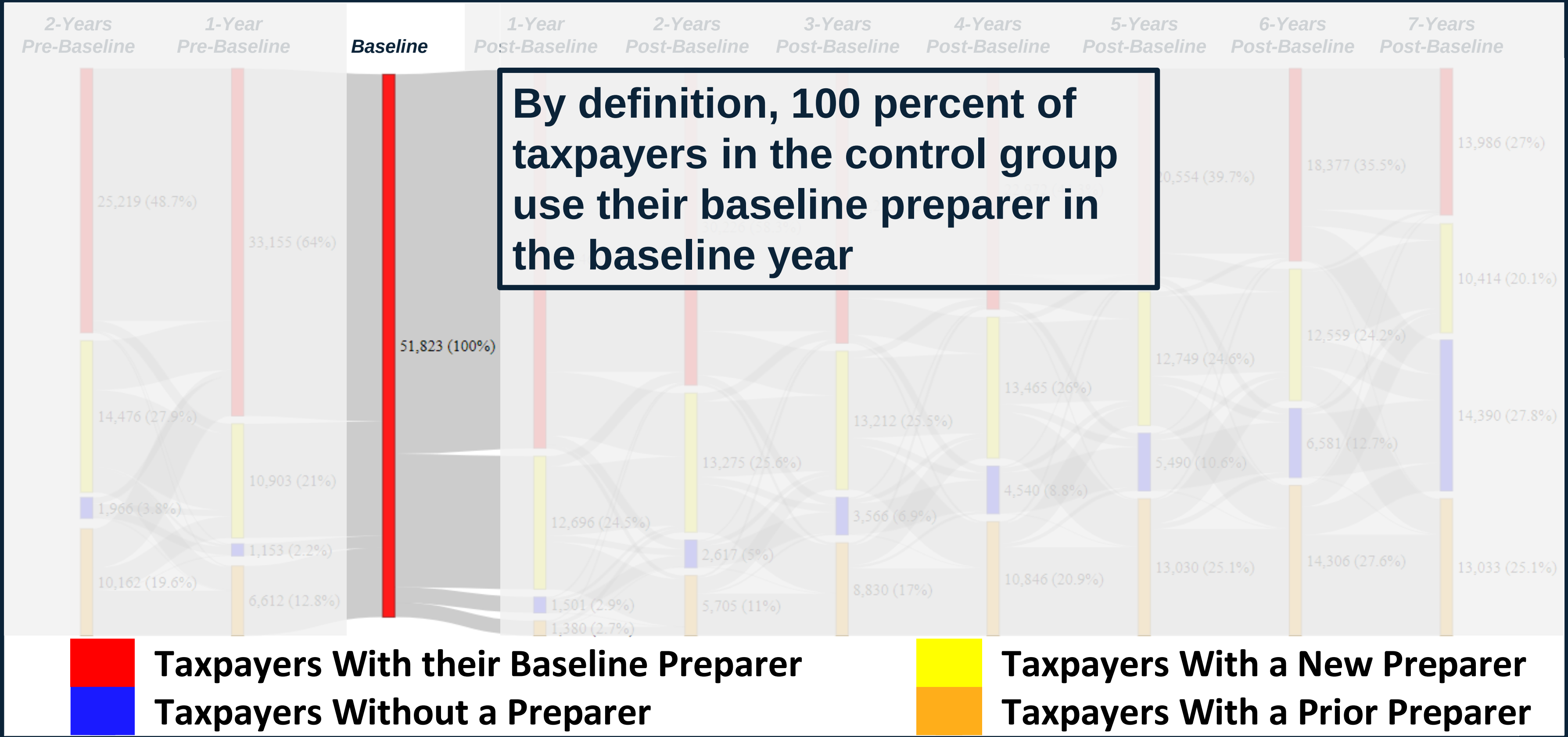
- **We use Sankey Diagrams to observe and analyze how taxpayers use their baseline year tax preparer over the study period**
- **Using tax filing data, we categorize each taxpayer for each year into the following discrete groups:**
  - Taxpayers with their Baseline Preparer
  - Taxpayers without a Preparer
  - Taxpayers with a New Preparer
  - Taxpayers with a Prior Preparer (not baseline preparer, but one used in years 1-7)
- **These groups allow us to compare client-preparer relationships between our treatment and control group taxpayers and identify potential behavioral differences that we theorize would affect the propagation of an indirect effect**

MITRE

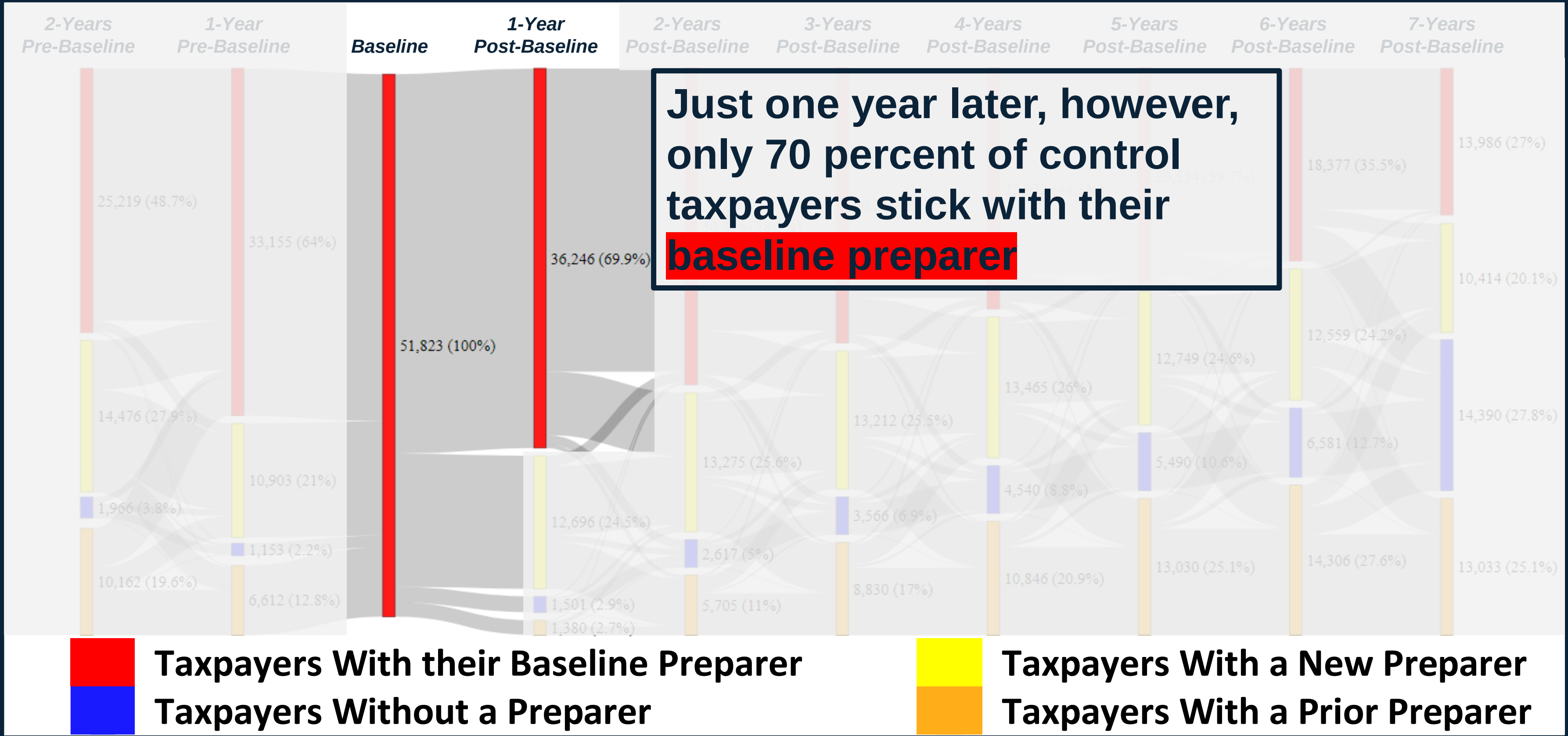
Taxpayer-Preparer Longevity among Control Group



Taxpayer-Preparer Longevity among Control Group

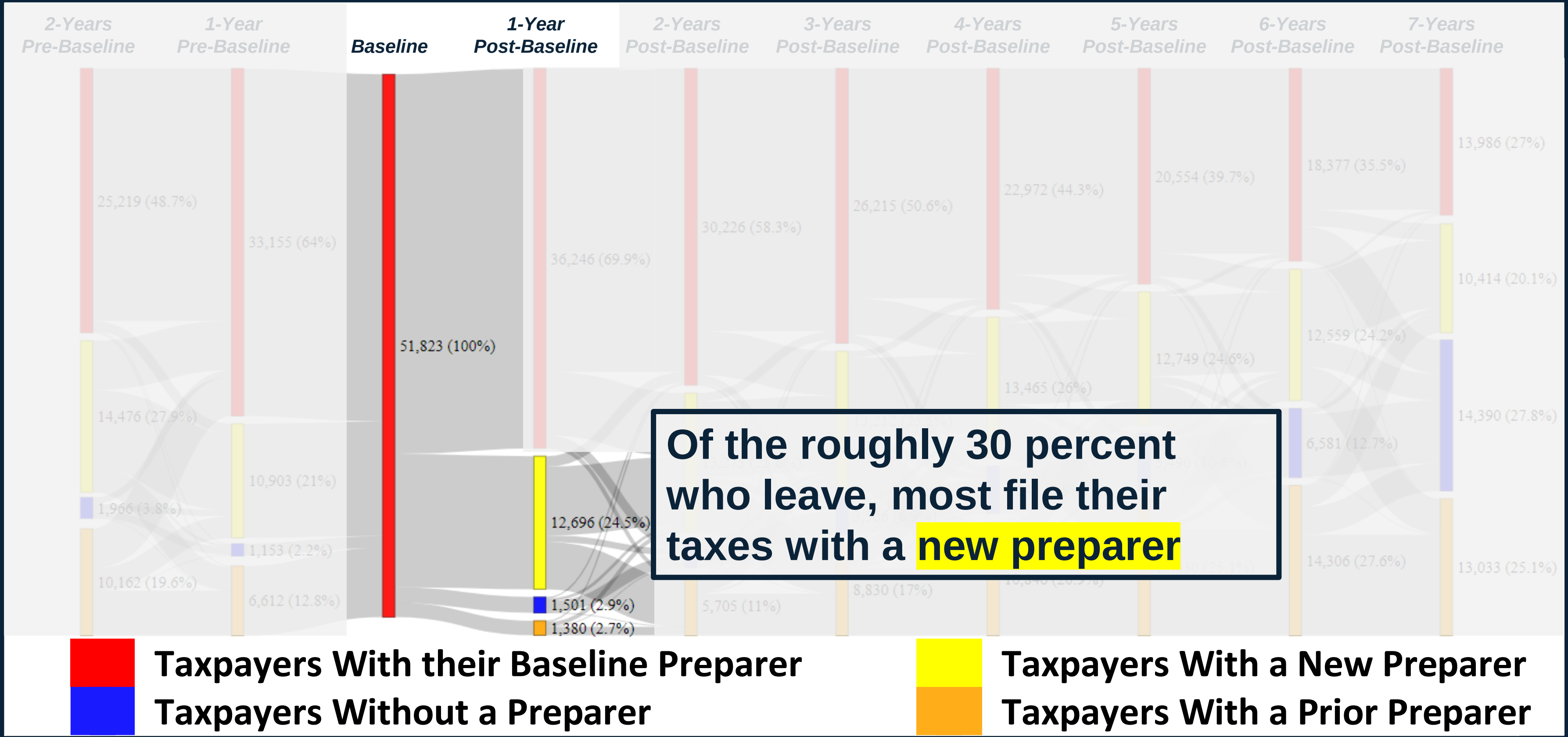


Taxpayer-Preparer Longevity among Control Group

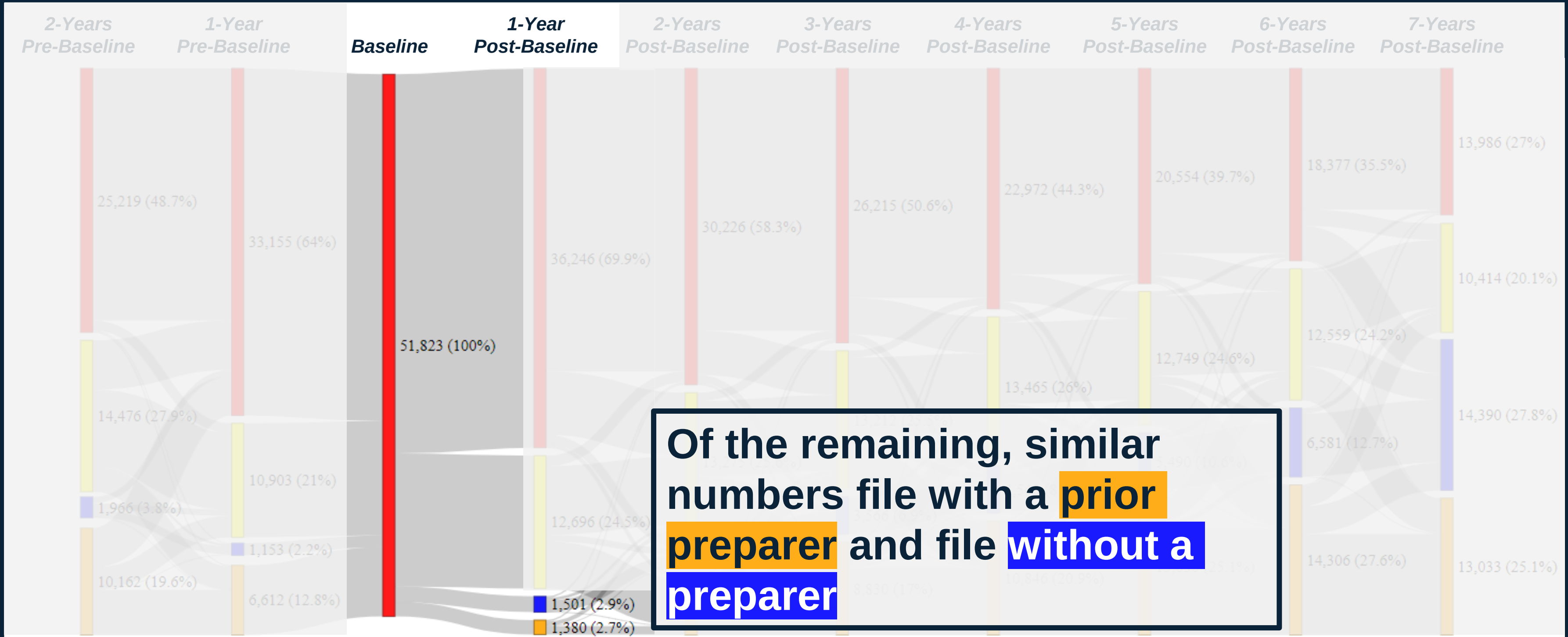


Just one year later, however, only 70 percent of control taxpayers stick with their baseline preparer

Taxpayer-Preparer Longevity among Control Group



Taxpayer-Preparer Longevity among Control Group



Of the remaining, similar numbers file with a prior preparer and file without a preparer

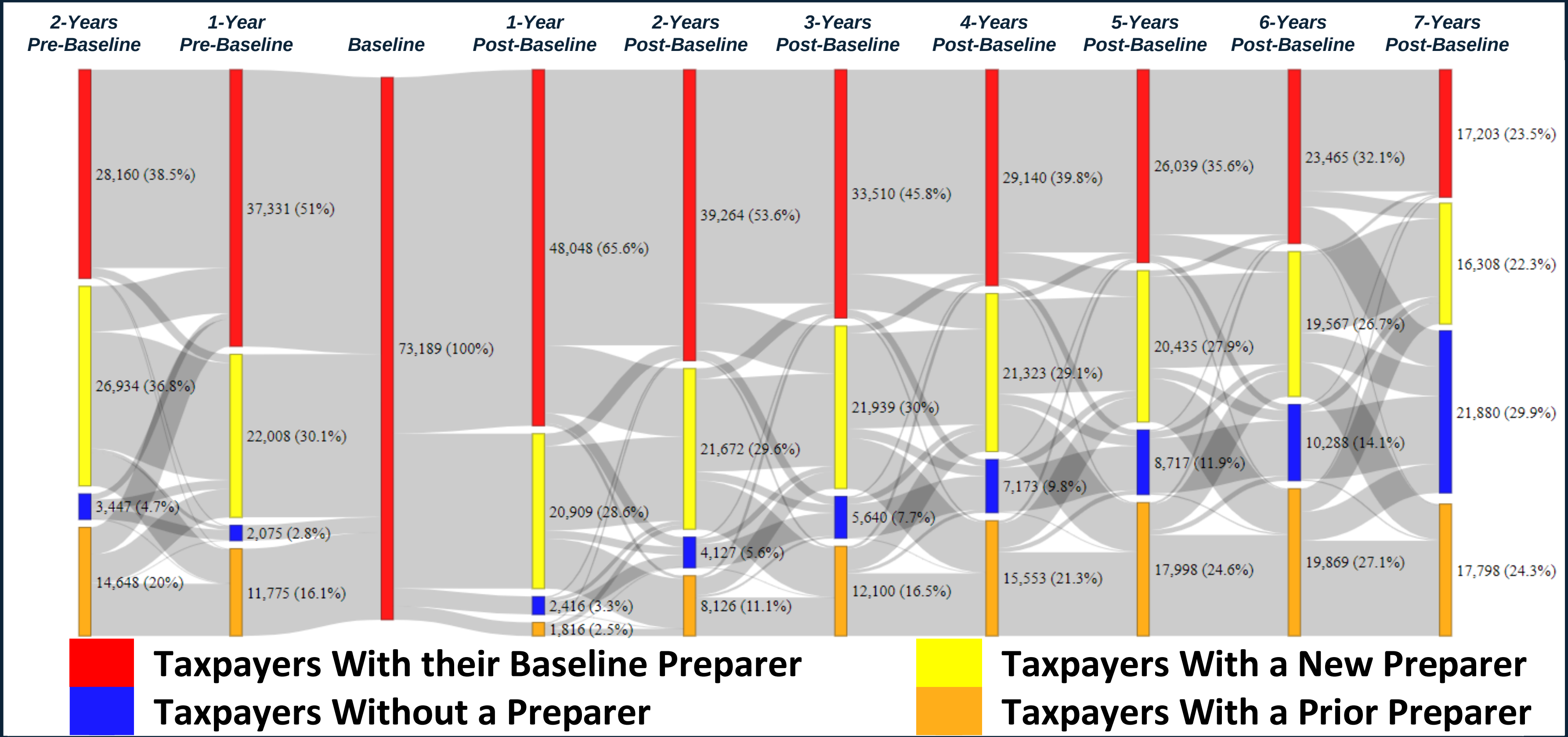
Taxpayers With their Baseline Preparer

Taxpayers Without a Preparer

Taxpayers With a New Preparer

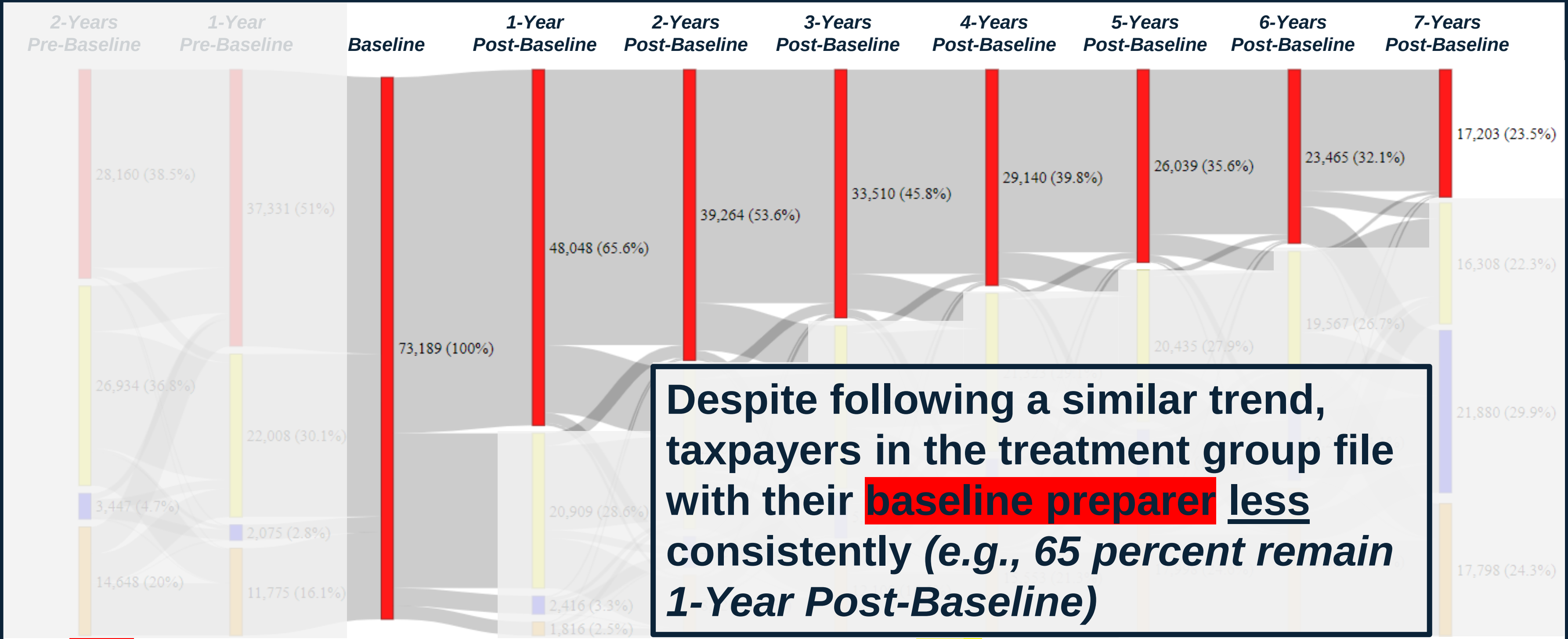
Taxpayers With a Prior Preparer

Taxpayer-Preparer Longevity among Treatment Group



MITRE

Taxpayer-Preparer Longevity among Treatment Group



Despite following a similar trend, taxpayers in the treatment group file with their baseline preparer less consistently *(e.g., 65 percent remain 1-Year Post-Baseline)*

Taxpayers With their Baseline Preparer
Taxpayers Without a Preparer
Taxpayers With a New Preparer
Taxpayers With a Prior Preparer

MITRE

Key Insights from Taxpayer-Preparer Sankey Diagrams

- In every year after the baseline, we see a steady decrease in taxpayers using their baseline year preparer.
- Taxpayer-Preparer relationships are somewhat **“sticky”**
    - That is, taxpayers who use their baseline preparer two years after the baseline are more likely to continue with their baseline preparer in the future.
- 25 percent of taxpayers in our sample employ a **new** tax preparer every year
- Taxpayers in our treatment group are slightly less likely to remain with their baseline tax preparer (~5 percent fewer relative to the control group)

These dynamics have real-world implications that must be considered when modeling

Linear Mixed Effects Model: *Total Tax*

For the i^{th} taxpayer, in the p^{th} preparer's network, in the j^{th} year after baseline:

$$\ln(\text{total tax} + 1)_{ipj}$$
$$= \beta_0 + \gamma_{0i} + \gamma_{0p} + \beta_1 treated_i + \beta_{2-9} year\ after\ baseline_{ij}$$
$$+$$

Where:

- Random intercepts for taxpayer (γ_{0i}) and tax preparer (γ_{0p}) included
- All dollar amounts are adjusted for inflation to 2018 USD

MITRE

Linear Mixed Effects Model: *Total Tax*

For the i^{th} taxpayer, in the p^{th} preparer's network, in the j^{th} year after baseline:

$$\ln(\text{total tax} + 1)_{ipj}$$
$$= \beta_0 + \gamma_{0i} + \gamma_{0p} + \beta_1 treated_i + \beta_{2-9} year\ after\ baseline_{ij}$$
$$+ \beta_{10-17} treated_i * year\ after\ baseline_{ij} +$$

Where:

- Random intercepts for taxpayer (γ_{0i}) and tax preparer (γ_{0p}) included
- All dollar amounts are adjusted for inflation to 2018 USD

Linear Mixed Effects Model: *Total Tax*

For the i^{th} taxpayer, in the p^{th} preparer's network, in the j^{th} year after baseline:

$$\ln(\text{total tax} + 1)_{ipj} = \beta_0 + \gamma_{0i} + \gamma_{0p} + \beta_1 treated_i + \beta_{2-9} year\ after\ baseline_{ij} + \beta_{10-17} treated_i * year\ after\ baseline_{ij} + \beta_{18} priority_i + \sum_{k=19}^{25} \beta_k C_k + \varepsilon_{ipj}$$

Where:

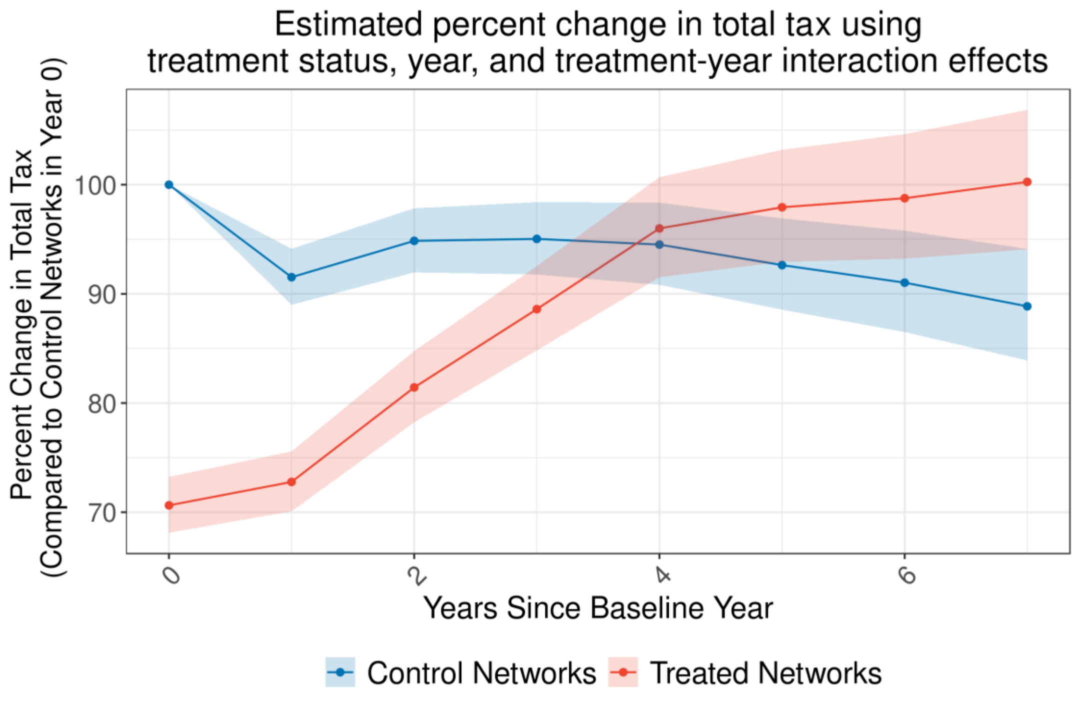
- Random intercepts for taxpayer (γ_{0i}) and tax preparer (γ_{0p}) included
- All dollar amounts are adjusted for inflation to 2018 USD
- C includes Tax Year, audit history, wage income, preparer's number of eligible clients, total positive income, and filing status

Results

TOTAL TAX MODELS

Results: Estimated Total Tax

- Stark difference in total tax reporting in baseline year
- Treated tax preparer networks see sharp increase in total tax reporting
- Control networks decrease slightly over time period
- However, this model does not capture the behavioral dynamics observed in the Sankey diagrams



Using Taxpayer-Preparer Dynamics in Total Tax Estimation

We know:

- There is a time delay of ~2 years between the filing year and audit start
- Roughly between 40-50 percent of taxpayers in our sample do not use their baseline preparer two years after the baseline year

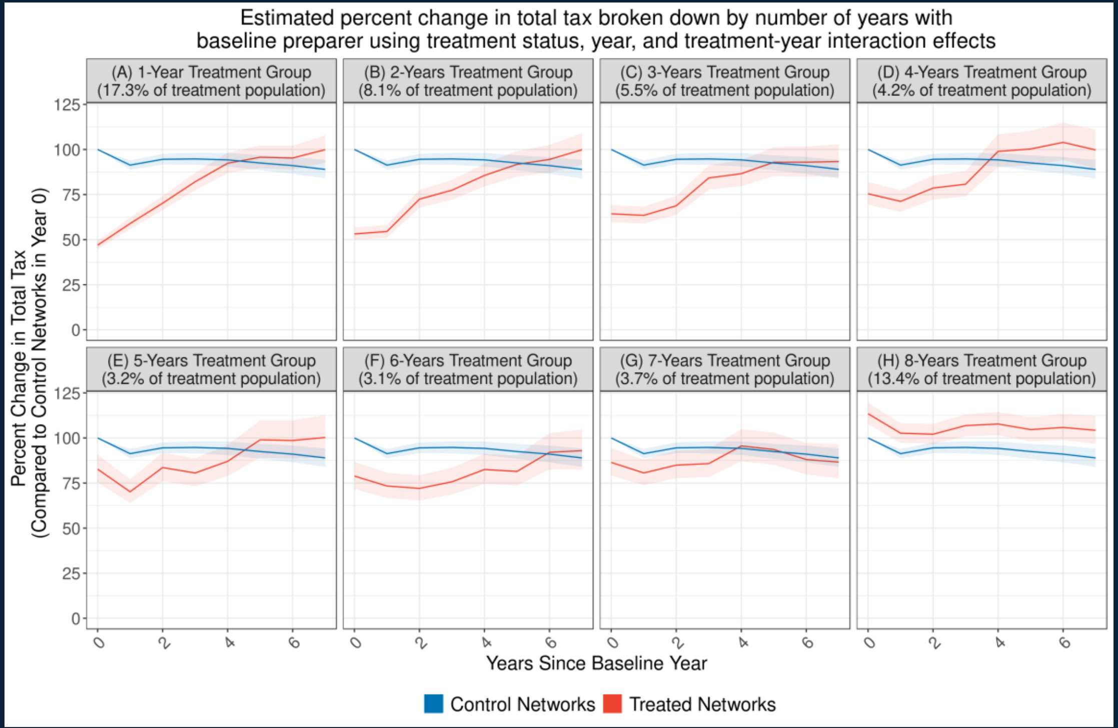
<u>Therefore, we should use this information to improve our model and estimation.</u>

To do so, we redefine our treatment variable as the *number of years* the taxpayer was with their baseline preparer

$$treated_i = \begin{cases} 0, & \text{if taxpayer } i \text{ is in the control group;} \\ \sum_{t=0}^{7} I(BaselinePreparer_i == TaxPreparer_{it}), & \text{if taxpayer } i \text{ is in the treatment group} \end{cases}$$

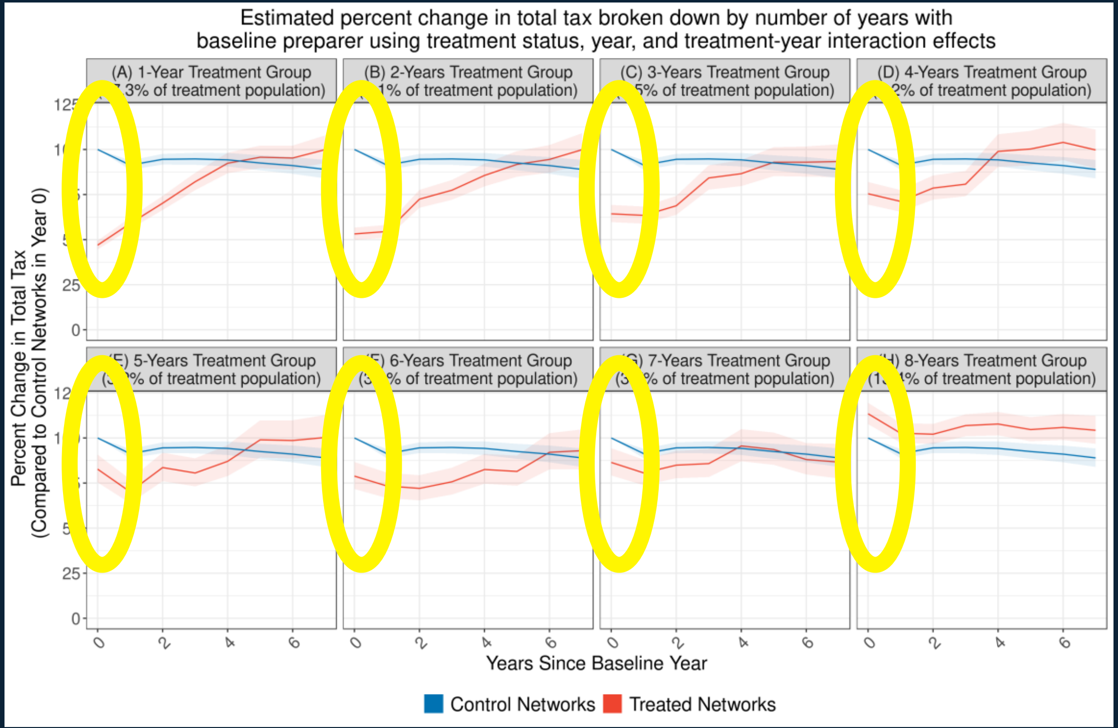
Results: Estimated Total Tax by Longevity

- Results show control group plotted against each treatment subset



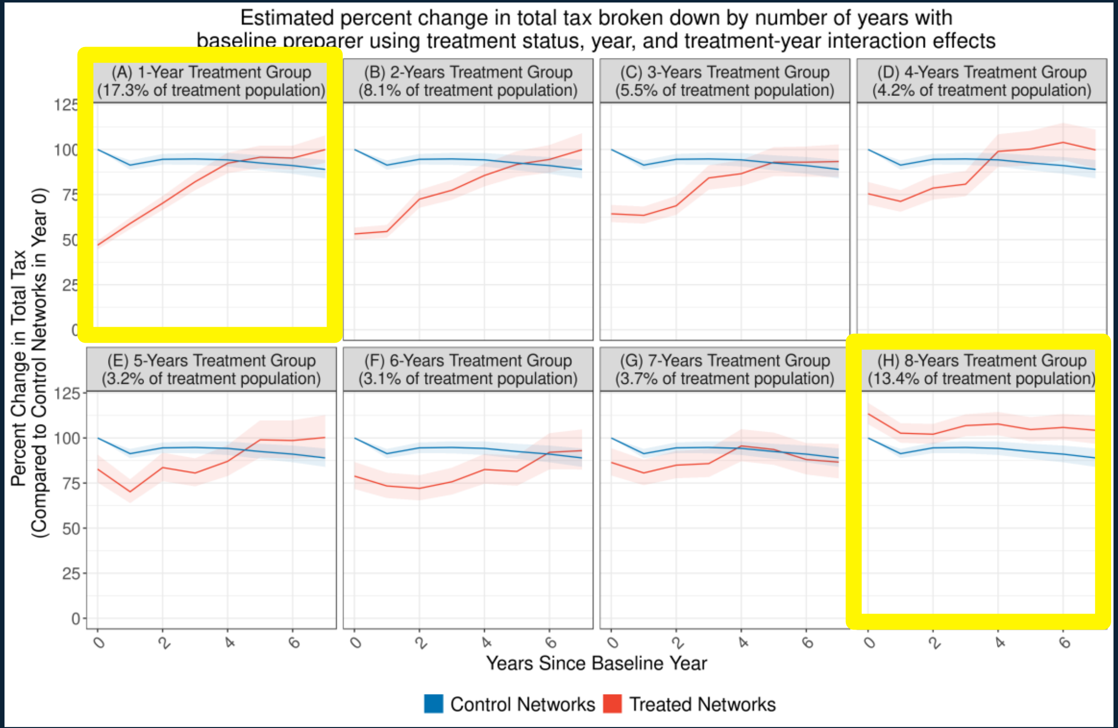
Results: Estimated Total Tax by Longevity

- Results show control group plotted against each treatment subset
- Baseline differences remain between these groups



Results: Estimated Total Tax by Longevity

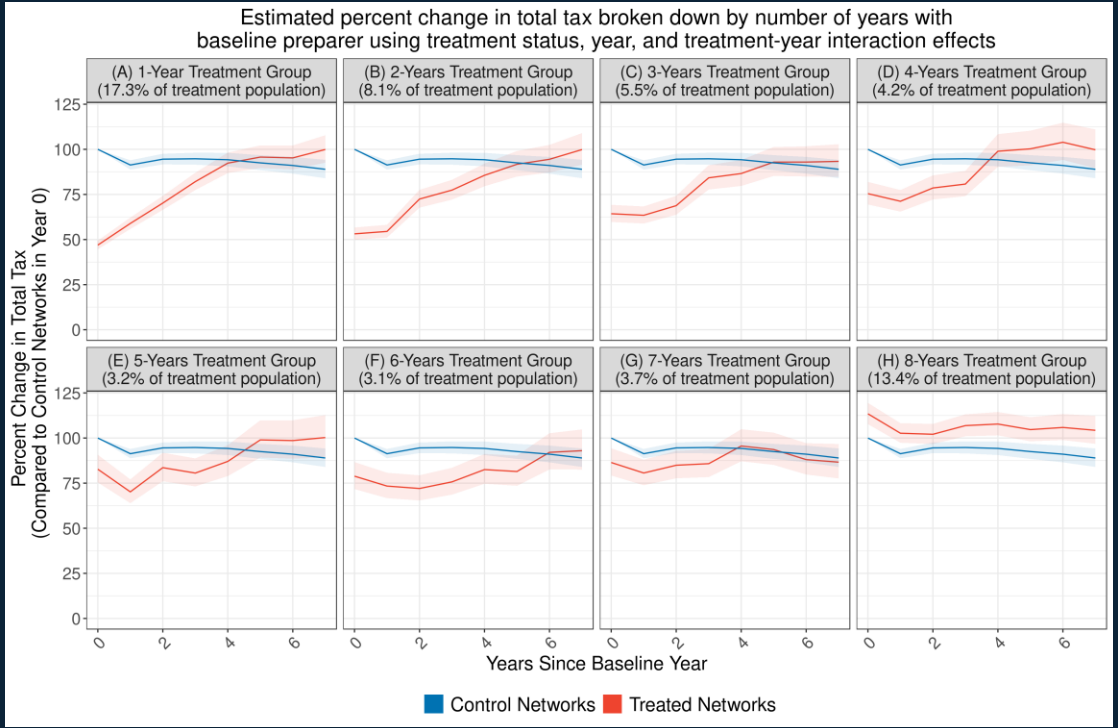
- Results show control group plotted against each treatment subset
- Baseline differences remain between these groups
- However, controlling for the duration of the taxpayer-preparer relationship yields starkly different estimates



MITRE

Results: Estimated Total Tax by Longevity

- Results show control group plotted against each treatment subset
- Baseline differences remain between these groups
- However, controlling for the duration of the taxpayer-preparer relationship yields starkly different estimates
- Overall, our findings suggest a **smaller** general indirect effect among taxpayers with longer preparer relationships



MITRE

Discussion

Discussion

- We find that about 30 percent of Sched. C audit-eligible taxpayers are with their baseline tax preparer for only one year
- Our findings are suggestive of a general indirect effect through tax preparer networks but are **counterintuitive** in some ways
  - When we disaggregated our findings by longevity of taxpayer-tax preparer relationship, we found that the effect was *less* strong for taxpayers who exhibited the longest duration with the tax preparer
- Our findings raise questions surrounding **selection bias vs. causality** in choosing a preparer.
  - Are taxpayers in higher audit likelihood situations seeking out specialized preparers?
  - Or…Are tax preparers encouraging higher business expense reporting which results in higher audit incidence in these networks?

Limitations

- Audit eligibility criteria may have shifted over time—we have access only to recent business rules
- Mechanism assumes audited taxpayers notify their baseline preparer
  - If they have moved on to a new preparer, their baseline preparer **may not be aware** of the audit at all
- Only considering one type of Schedule C audit within the preparer network

Future Research

- Quantify the general indirect effect **in dollars** such that it can be used to inform resource allocation
- Explore **alternative definitions** of the “baseline” year to account for the lag in audit notification

Contact us:

Kyle Furlong krfurlong@mitre.org
Leigh Nicholl lnicholl@mitre.org
Ellen Badgley ebadgley@mitre.org
Lucia Lykke, Ph.D. llykke@mitre.org
Alan Plumley, Ph.D. Alan.Plumley@irs.gov

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD™

Do Collateral Sanctions Work? Evidence from the IRS' Passport Certification and Revocation Process

Paul R. Organ
Michigan

Alex Ruda
IRS

Joel Slemrod
Michigan

Alex Turk
IRS

June 24, 2021

IRS-TPC Joint Research Conference

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

Agenda

- Background and motivation
- Direct effects of the program
  - Passport application denials
  - RCT analysis of certification
- Policy implications
  - Marginal revenue estimates

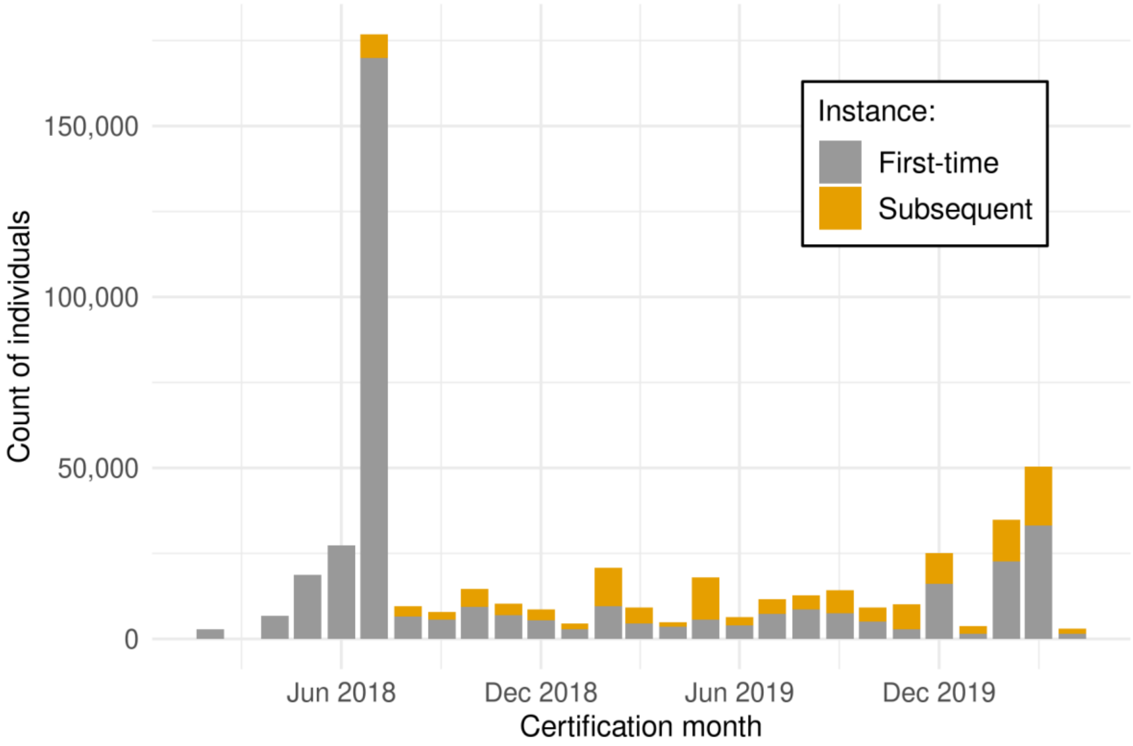
IRS recently implemented passport sanctions

- 2015's FAST Act required IRS and State Dept. to implement a new collateral sanction: **restricting passport access for "seriously delinquent" tax debtors**
- Eligibility for passport **certification**:
  - Total eligible debt > $50,000 (now $54,000)
  - Already attempted a Lien or Levy:
    - Notice of Federal Tax Lien filed and Collection Due Process hearing rights expired, or
    - Notice of levy issued
  - Not excluded due to hardship, bankruptcy, in an Installment Agreement, etc.
- Consequences:
  - New passport applications will not be approved
  - Cannot make changes to existing passport, or renew
  - In certain cases, can have existing passport revoked (IRS request, State Dept. discretion)

Why consider collateral sanctions?

- **Collateral sanction**: applied in addition to formal tax penalties; rescinds government-provided benefit or privilege; usually enforced by non-tax agency
- May be more effective than monetary penalties (Blank (2013))
  - Salience; loss aversion; reputation effects
- Kuchumova (2018) formalizes rationale for their use
  - Collateral sanctions affect consumption and allow targeted enforcement
  - Can allow gov't to impose punishment correlated with earning potential
- Other examples of collateral sanctions for tax purposes
  - Federal: FHA mortgage eligibility; contracting with federal government
  - States: drivers' licenses and vehicle registrations (CA); law and professional licenses (MN, WI); hunting and gaming permits (LA)

Large initial rollout, then steady flow of certifications

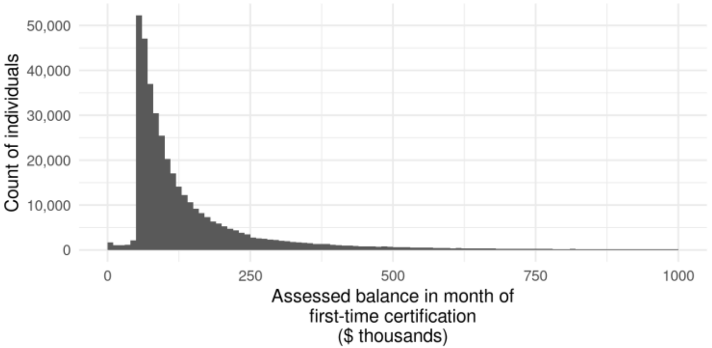


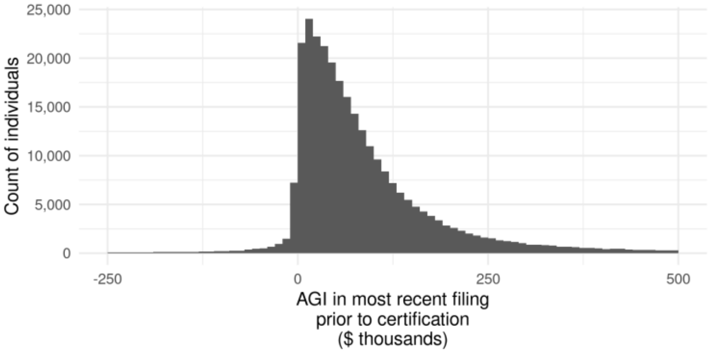
Note: new certifications stopped in April 2020 due to Covid-19

First-time certifications: summary statistics

	Mean	St Dev	25th Pctile	Median	75th Pctile	# Obs
Certified balance						
Assessed balance, penalties, and interest ($ thousands)	$\boxed{\$197}$	$1,147	$68	$\boxed{\$98}$	$172	393,000
Number of modules	5	4	2	4	7	393,000
Age of oldest module (years)	7	3	4	7	9	393,000
Most recent tax filing prior to certification						
Total positive income ($ thousands)	$149	$5,549	$30	$68	$134	293,000
Adjusted gross income ($ thousands)	$\boxed{\$103}$	$1,925	$24	$\boxed{\$60}$	$120	293,000
Age in 2017 (years)	53	11	46	53	61	379,000

Note: Rounded for disclosure purposes; First-time certifications through April 2020

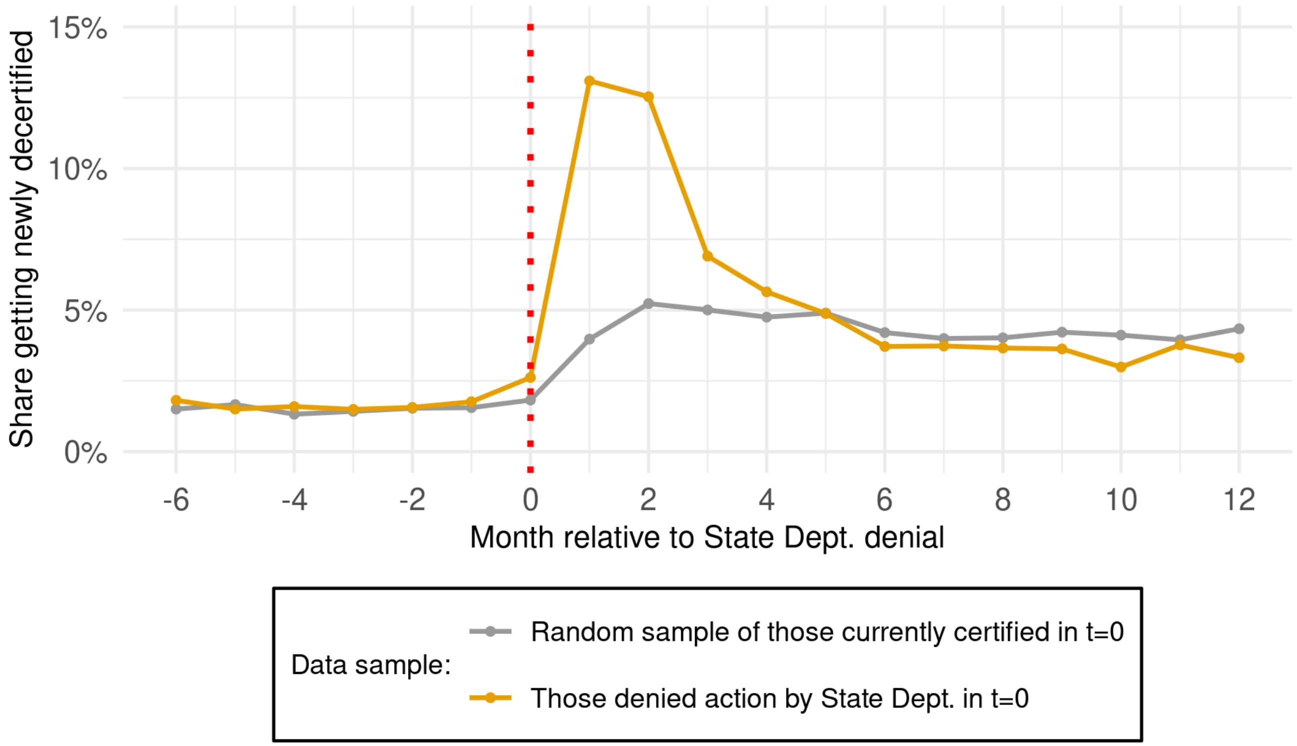




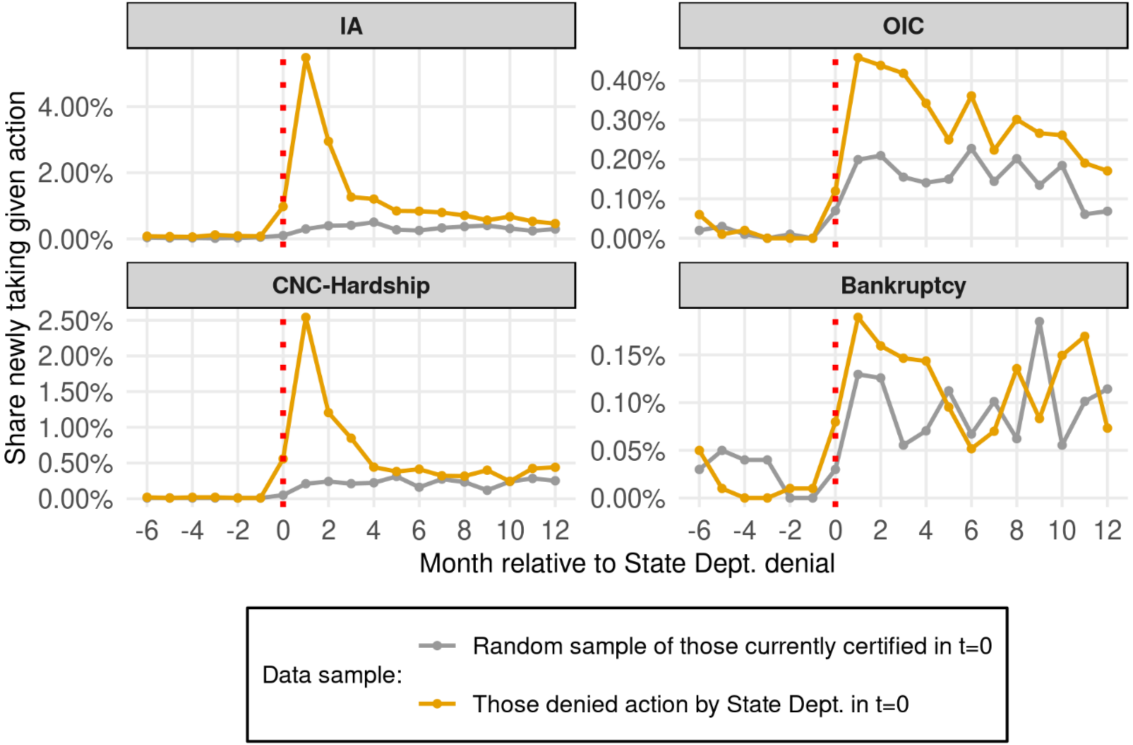
Passport request denials by State Dept.

- Through Jan 2020, about 12K taxpayers made requests to the State Dept. and were denied at least once due to certification
  - Includes new applications, renewals, and modifications
  - We study 10K of these: first-time denials, grouping joint liabilities
- Clearest opportunity to study effect of certification – this is a group that are clearly treated (i.e., they clearly have or want passports)
- Graphical approach – observe behavior before and after denial
  - Potential concern: "mechanical effect" (conditioning on being certified in t0 means not taking action in prior months, and some probability will take action following months)
  - Solution: randomly select a control group of certified taxpayers, conditioning on same thing (certified in time t=0) to identify mechanical effect

Denied requests lead to immediate action



Denied requests lead to immediate action

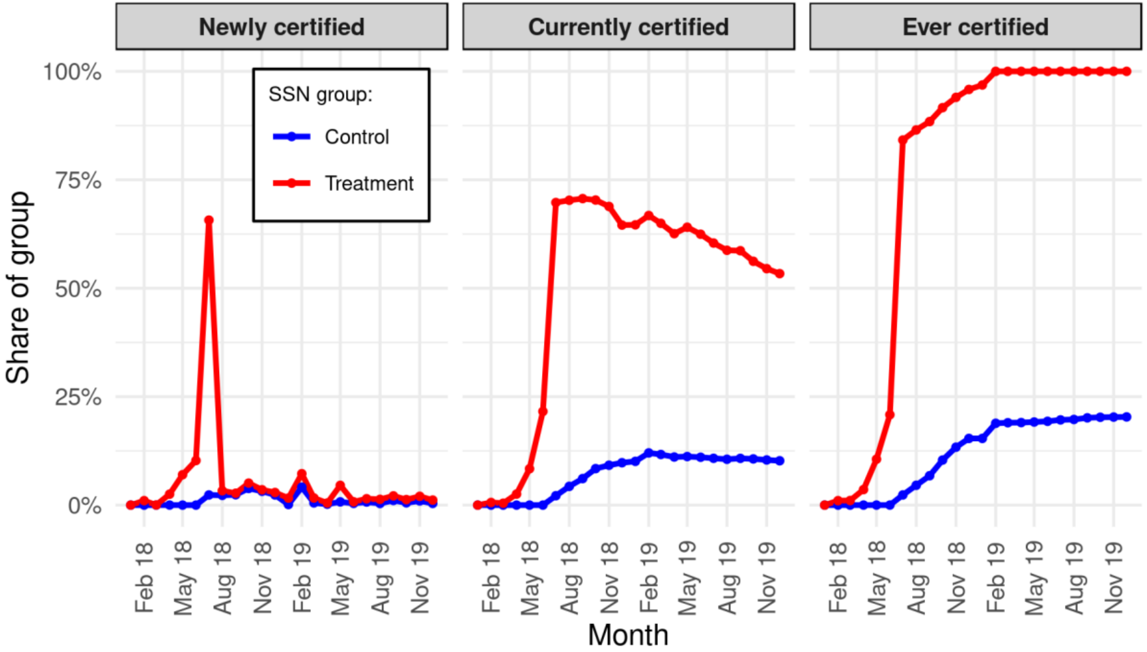


Denied requests lead to immediate action

Effect of certification – RCT approach

- Use RCT-style rollout to test direct effects of certification
- Context:
  - During rollout of passport program (March 2018-February 2019)
  - SSN used to “hold out” 5% of eligible taxpayers
  - First time taxpayer came up as eligible, if in holdout group, not certified
  - Caveat: some of these held out taxpayers came up for eligibility again, due to certain actions or changes in account characteristics, and were then certified
- We will use an IV approach to recover the direct effect of certification
  - Similar to the MDVE discussion in Angrist (2005) (“treatment migration”)
  - SSN randomization is “intent to treat”
- Note: control taxpayers may have known they were eligible but not yet certified, and could have taken action to try to avoid certification => our results are a lower bound on effect of certification

Timing and prevalence of certifications



Note: taxpayers here are split only by their SSN, i.e., by intent-to-treat. 100% of the treatment SSN group (in red) get certified during the RCT phase, while only 19% of the control SSN group (in blue) get certified during the RCT phase.

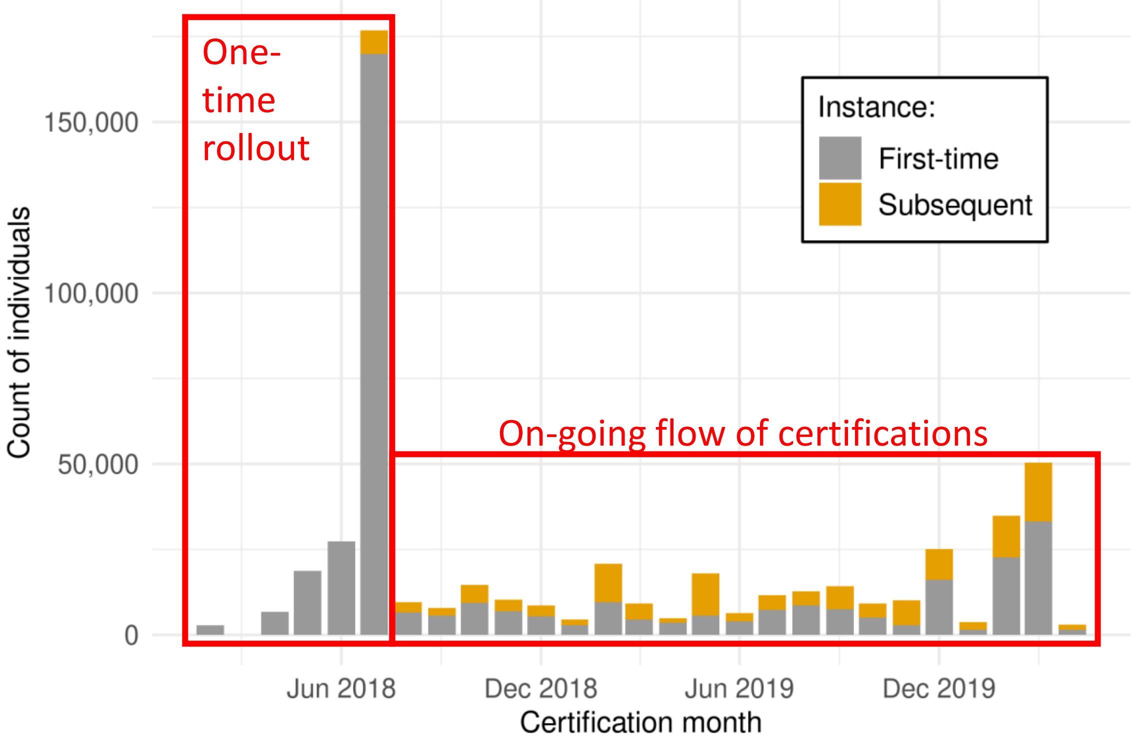
RCT: IV regression approach

- Regression specifications:
  - Structural equation: $Outcome_i = \alpha Cert_i + X_i'\beta + \epsilon_i$
  - First stage: $Cert_i = \pi_0 SSN_i + X_i'\pi_1 + \eta_i$
  - Reduced form: $Outcome_i = \delta_0 SSN_i + X_i'\delta_1 + \epsilon_i$
  - IV coefficient of interest: $\alpha = \delta_0/\pi_0$
- Data sample:
  - Taxpayers certified or held out during RCT phase (Mar '18 – Feb '19)
  - Restrictions based on Dec' 17 balances:
    - Total assessed balance <$1M
    - Max module age < 12 years
    - Number of modules < 10 (annual) and <40 (quarterly)

Certification causes new compliance actions

	Taking new action any time Mar '18-Dec '19				Fully resolved as of Dec '19		Any of six listed actions
	IA	OIC	CNC	Bankruptcy	By payment	By abatement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Certified	**0.0131*** **	**0.0029***	**0.0049** **	**-0.0004**	**0.0016**	**0.001**	**0.0205*** **
	(0.0026)	(0.0017)	(0.0020)	(0.0014)	(0.0013)	(0.0009)	(0.0037)
… (see paper for full table)							
Observations	266,890	266,890	266,890	266,890	266,890	266,890	266,890
Adjusted R^2	0.075	0.022	0.020	0.010	0.021	0.012	0.124
Mean dep. var.	0.076	0.028	0.040	0.017	0.016	0.008	0.172
Effect as percent of mean:	17%	10%	12%	-2%	10%	13%	12%

Pre-existing cases vs. new cases



Note: new certifications stopped in April 2020 due to Covid-19

Marginal revenue estimates

		Pre-existing cases	New cases
Effect of denied requests			
Estimated payment effect	*\$/person*	\$7,300	\$12,600
Effect of certification			
Coefficient estimate	*% of balance*	0.11%	0.57%
Average total balance	*\$*	\$202,000	\$195,000
Estimated payment effect	*\$/person*	\$222	\$1,112

Conclusions

- Passport program is an example of an effective collateral sanction
  - Strong response among those denied a passport-related request
    - Includes some full pay resolutions, as well as other actions (IAs, OICs, CNCs)
  - Certification leads to more compliance actions, especially IAs (17% increase)
- See paper for more:
  - Heterogeneity analyses
  - Robustness checks
  - Welfare discussion
- Comments welcome! **prorgan@umich.edu**

EITC Noncompliance: Examining the Roles of the Dynamics of EITC Claims and Paid Preparer Use

Emily Y. Lin, Ankur Patel, and Alexander Yuskavage

Office of Tax Analysis

U.S. Department of the Treasury

Introduction

- Paper goals:
  - Plot out individual trends in EITC claims and preparation methods
  - Understand what drives those trends
  - Understand how those trends relate to compliance
- We identify several important trends
  - Shift towards uncredentialed preparers
  - Multiple types of EITC ‘churn’
- We find that past claim outcome affects current behavior

Motivation

- EITC compliance has remained stable over past ~20 years
- EITC claims are considered complex, especially considering participant income levels
- ‘Knowledge base’ of EITC incorporates paid preparers and past participants

Lit review

- Relationship between tax preparers and compliance is complicated
  - Klepper and Nagin (1989), Lin (2020), DeBacker et al (2021)
- Tax Preparer skill matters…
  - Langetieg et al (2013), GAO (2006), TIGTA (2004)
- …but individual circumstances matter too
  - Saez (2010), Chetty et al (2013)

Data

- IRS administrative data from 2010-2018
- Panel constructed for 2% sample of 2010 EITC claimants with qualifying children
- Do not have direct compliance measures for most observations
  - We use two different indirect compliance measures

Historic trends in EITC claims

Shares of New Claimants and Return Preparation Methods by Year					
Year	Claimants (million)	New Claim	E-File	Use Any Preparer	Unknown or No Credential
2003	17.71	24%	14%	75%	
2004	17.82	25%	16%	75%	
2005	18.04	25%	17%	76%	
2006	18.24	25%	19%	75%	
2007	18.79	25%	21%	74%	
2008	19.13	25%	24%	72%	
2009	20.80	27%	27%	69%	
2010	20.86	23%	30%	67%	50%
2011	20.99	23%	33%	65%	48%
2012	20.90	22%	36%	62%	46%
2013	21.19	23%	38%	60%	44%
2014	20.88	21%	40%	59%	43%
2015	20.66	21%	42%	57%	42%
2016	20.20	20%	42%	56%	42%
2017	19.72	20%	42%	56%	42%
2018	19.16	20%	43%	55%	41%

Transitions between Claim Types

Transition of EITC Status 2011-2018 for Claimants in 2010				
Tax Year	EITC Claim in Year t-1		No EITC Claim or Not Filing in Year t-1	
	Claim in t	Stop Claim in t	Resume Claim in t	Continue No Claim in t
2011	0.779	0.221	.	.
2012	0.634	0.145	0.057	0.164
2013	0.569	0.121	0.063	0.246
2014	0.517	0.116	0.060	0.308
2015	0.469	0.108	0.059	0.364
2016	0.425	0.103	0.056	0.416
2017	0.385	0.096	0.054	0.466
2018	0.349	0.090	0.052	0.510

Preparation Methods

- Different classes of preparation assistance
  - Credentialed Preparers
  - Uncredentialed Prepared
  - E-filing
- Partial data on credentialed preparers starting in 2010

Preparation Method Transitions

Shares of Return Preparation Methods 2011-2018 for 2010 Claimants					
Tax Year	Credentialed Preparer	Unregulated Preparer	VITA/TCE	Self-Prepared	Sample Size
2010	1.000				62,726
2011	0.751	0.184	0.004	0.061	46,722
2012	0.667	0.224	0.007	0.102	40,330
2013	0.612	0.249	0.009	0.130	36,020
2014	0.568	0.264	0.010	0.158	32,097
2015	0.534	0.275	0.010	0.181	28,683
2016	0.505	0.286	0.011	0.198	25,487
2017	0.474	0.303	0.011	0.213	22,710
2018	0.448	0.310	0.010	0.232	20,146

Preparation Method Determinants

Use of a Paid Preparer				
	OLS, 2010-2018	FE LPM, 2010-2018	FE LPM, 2013-2018	FE LPM, 2015-2018
Preparer Usage, Prior Year	0.664***	0.231***	0.139***	-0.030***
	(0.001)	(0.001)	(0.001)	(0.002)
Log # Prior Preparer Uses	0.103***	0.345***	0.516***	0.860***
	(0.000)	(0.002)	(0.003)	(0.006)
EIC Claim, Prior Year	0.021***	0.022***	0.020***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)
Log # Prior EITC Claims	-0.049***	-0.140***	-0.138***	-0.145***
	(0.000)	(0.002)	(0.003)	(0.006)
N	2,364,313	2,364,313	1,701,694	1,100,778

Indirect compliance flags

- Dependent Database (DDb) score
  - Contains decision rules for EITC criteria
  - We focus on EITC returns with qualifying children
  - Having at least one DDb violation treated as a flag

- Discriminant Function (DIF) score
  - Used as part of audit determination process
  - Scores ranked by activity code
  - DIF score in top 80^{th} or 90^{th} percentile treated as flags

Compliance Flag validation

Correlation Between Indicators of Likely Noncompliance and Actual Noncompliance		
	Post-Audit EITC	Post-Audit Total Tax
Reported EITC	0.742***	
	(0.000)	
Reported Total Tax		1.589***
		(0.000)
DDB Flag	-638.624***	70.550***
	(0.272)	(0.692)
DIF10 Flag	-191.596***	570.928***
	(0.491)	(1.261)
DIF20 Flag	-45.933***	320.068***
	(0.364)	(0.804)
N	136,919,373	136,919,373

Compliance

- Unclear how information spreads via prepared returns
  - Preparers provide ‘example’ for future returns
  - Clients of preparers are self-selecting on risk, knowledge, etc
- Unclear how information spreads via experience
  - Better understanding of how to fill out forms…
  - …can be used for good or for ill

Past Compliance

Past EITC Claiming Experience and Current Claim Noncompliance					
Number of Prior Claims With Flag	Violates Any DDb Rule	DIF Score in Top Decile	**Number of Prior Claims Without Flag**	Violates Any DDb Rule	DIF Score in Top Decile
1	-0.06	-0.026	**1**	0.115	0.079
	(0.001)	(0.000)		(0.001)	(0.001)
2	-0.095	-0.034	**2**	0.189	0.111
	(0.001)	(0.001)		(0.001)	(0.001)
3	-0.114	-0.038	**3**	0.226	0.145
	(0.001)	(0.001)		(0.002)	(0.002)
4+	-0.095	-0.044	**4+**	0.189	0.272
	(0.001)	(0.001)		(0.003)	(0.004)

Compliance and Preparers

Past Preparer Use and Current Claim Noncompliance					
Type of Prior Claims With Flag	Violates Any DDb Rule	DIF Score in Top Decile	**Type of Prior Claims Without Flag**	Violates Any DDb Rule	DIF Score in Top Decile
Self-Prepared	0.000	0.000	**Self-Prepared**	0.029	-0.033
	(0.001)	(0.001)		(0.003)	(0.003)
Credentialed Preparer	0.002	0.003	**Credentialed Preparer**	-0.02	0.001
	(0.001)	(0.001)		(0.003)	(0.003)
Uncredentialed Preparer	0.000	-0.001	**Uncredentialed Preparer**	0.031	-0.043
	(0.001)	(0.001)		(0.003)	(0.003)

Conclusions

- There has been a large shift away from use of credentialed preparers
- Preparer credentials affect the degree and type of noncompliance
- Current compliance depends strongly on past compliance
- So far, all findings are correlation only
  - Next step is to address causality with IV, policy changes, etc

Discussion for Improving Individual Taxpayer Compliance

IRS-TPC Conference on Tax Administration

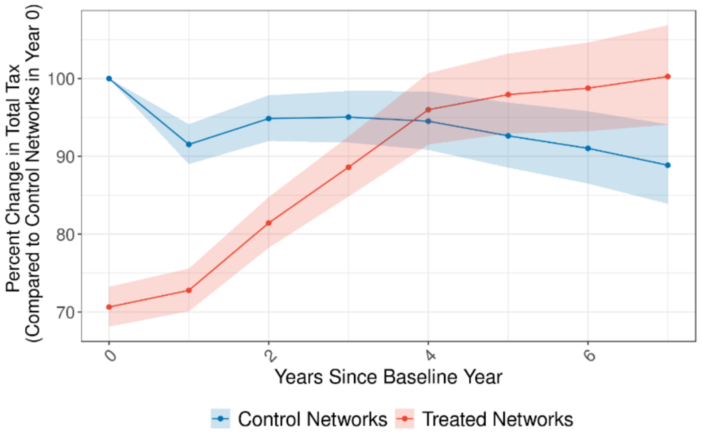
June 24, 2021

Tatiana Homonoff, NYU-Wagner

Audit Contagion? Furlong, Badgley, Lykke, Nicholl, and Plumley

- There is a large literature estimating the direct effect of audits, but much less work on indirect/spillover effect
  - Small literature on geographic network spillovers
  - Newer literature posits tax preparers as important network (Boning et al; 2020)

- This paper: compare audit-eligible taxpayers who did **not** experience an audit themselves, but did/did not have an audit in their tax preparer network



Identification Strategy

- Identification strategy: less selection than if comparing audited to audit-eligible taxpayers, but there still might be differences in taxpayers using treatment vs. control preparers
  - e.g., taxpayers in the treated network have higher audit priority, remit 29% less tax in the baseline year, less likely to remain with the baseline preparer, etc.
  - Analysis yields a puzzling result: the effect is strongest among those who remain with their preparer for only one year (i.e., switch **before** the audit takes place)
- Are there other restrictions or matching strategies to create a more balanced comparison?
- Alternatively, could you use the timing of the audit as variation and restrict the sample to only treatment preparers?

Mechanisms

- The authors make a reasonable assumption that tax preparers are informed about the audit and imply that the behavior change is driven by the preparers
    - Inclusion of an objective function for tax preparers (vs. individuals) would be helpful
    - e.g., how do we think about something like Allingham-Sandmo model here?

- Could the effects be driven by behavior changes among non-audited taxpayers who learn about the in-network audit?
    - Related: if so, do tax preparers need to worry about reputational concerns in their objective function?

- Also related: if taxpayers learn about audits in their preparer network, are they less likely to stay with their current preparer? Testable…
    - Is this related to the observed difference in effects by taxpayer-preparer longevity?

Do Collateral Sanctions Work? Organ, Ruda, Slemrod, and Turk

- This paper estimates the effects of a recent collateral sanction program: revocation of passport access for those with substantial tax debt

- Action packed paper! Three different analyses:
    1. RCT during rollout of the program (direct effect)
    2. Event study following passport denials (direct)
    3. Bunching analysis around $50,000 at program initiation (indirect)

- Results:
    - Small but significant effects on full population of certified taxpayers and larger immediate effects for those seeking (and then denied) passport access
    - Inconclusive indirect effect of program

Target Population Characteristics

- To be eligible for certification (i.e., rescinded passport access)
  - Substantial debt amount ($54,000 in 2021) and
  - A Notice of Federal Tax Lien filed and the associated Collection Due Process (CDP) hearing rights have expired, or a Notice of Levy has been issued

- Who exactly are the individuals who make it to this stage? How do they compare to populations in studies of more traditional compliance instruments (e.g., fines)?

- Note: all the more surprising that you observe large effects!

Other Comments

- Difficult to interpret heterogeneity by income: proxy for passport demand or ability to pay?
  - Could you do a similar heterogeneity analysis for your “State Denied” identification strategy to disentangle the two?

- Indirect effects: Bunching at debt threshold occurs prior to program implementation, possibly due to streamlined Installment Agreement (IA) policies using the same threshold
  - Could difference out pre-period years?
  - Also, streamlined IA process seems worthy of study since it seems to lead to substantial bunching!

EITC Noncompliance Lin, Patel, and Yuskavage

- This paper examines the relationship between EITC noncompliance and tax preparation method as well as prior noncompliance

- Finds that noncompliance is highest among:
  - New claimants
  - Users of unregulated preparers (the most common method in most years)
  - Prior non-compliers

- …and lowest among:
  - VITA/TCE users
  - Prior compliers

Tax Preparation Findings

- (Selfishly for my own work) I found the results on compliance by preparer type really interesting!

- Goldin, Homonoff, Javaid, & Schafer (2021): analyze an IRS experiment which provides information about free tax prep methods (VITA/TCE and Free File) to prior nonfilers
  - Finds significant increases in filing and, in turn, in EITC claiming
  - The current papers suggest that these new claims are also potentially more likely to be in compliance, especially if using VITA/TCE

- Interested in seeing this analysis with controls for taxpayers/return characteristics (i.e., combining Table 3 & 6)

Other Comments

- Many of the findings are characterized as evidence of learning
  - But are these findings not simply evidence of persistence?
  - e.g., the results can simply be the result of taking the same action every year rather than learning to be compliant or, conversely, to game the audit process

- How should we think about instances where findings for the two measures of noncompliance (DDb violations vs. DIF Score) differ?
  - e.g., credentialed preparers are less likely to have DDb violations but more likely to have high DIF scores

- Outcomes focus on noncompliant claims, but is there persistence in *any* claiming by prior compliance/preparation method?

Session 1. Improving Individual Taxpayer Compliance

Moderator:	***Robert McClelland*** *Tax Policy Center*
Audit Contagion? Investigating the General Indirect Effect of Audits Through Tax Preparer Networks	***Kyle Furlong*** *MITRE Corporation*
Do Collateral Sanctions Work? Evidence from the IRS' Passport Certification and Revocation Process	***Paul Organ*** *University of Michigan*
EITC Noncompliance: Examining the Roles of the Dynamics of EITC Claims and Paid Preparer Use	***Alexander Yuskavage*** *Treasury Office of Tax Analysis*
Discussant:	***Tatiana Homonoff*** *New York University*

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

Session 2. Impacts of Variations in Process

Moderator:	***Mary-Helen Risler*** *IRS: RAAS*
Sales Tax Administration and the Real Economy	***Roger White*** *Arizona State University*
Using Discrete Event Simulations to Understand the Impact of Changes to IRS Processes	***Rafael Dacal*** *IRS: SB/SE*
Effects of Post-filing Adjustments on Statistics of Income (SOI) Estimates	***Chloe Gagin*** *IRS: RAAS*
Discussant:	***Jennifer Stratton*** *Government Accountability Office*

Local Sales Tax Administration and the Real Economy

Jenny Brown
David Kenchington
Roger White

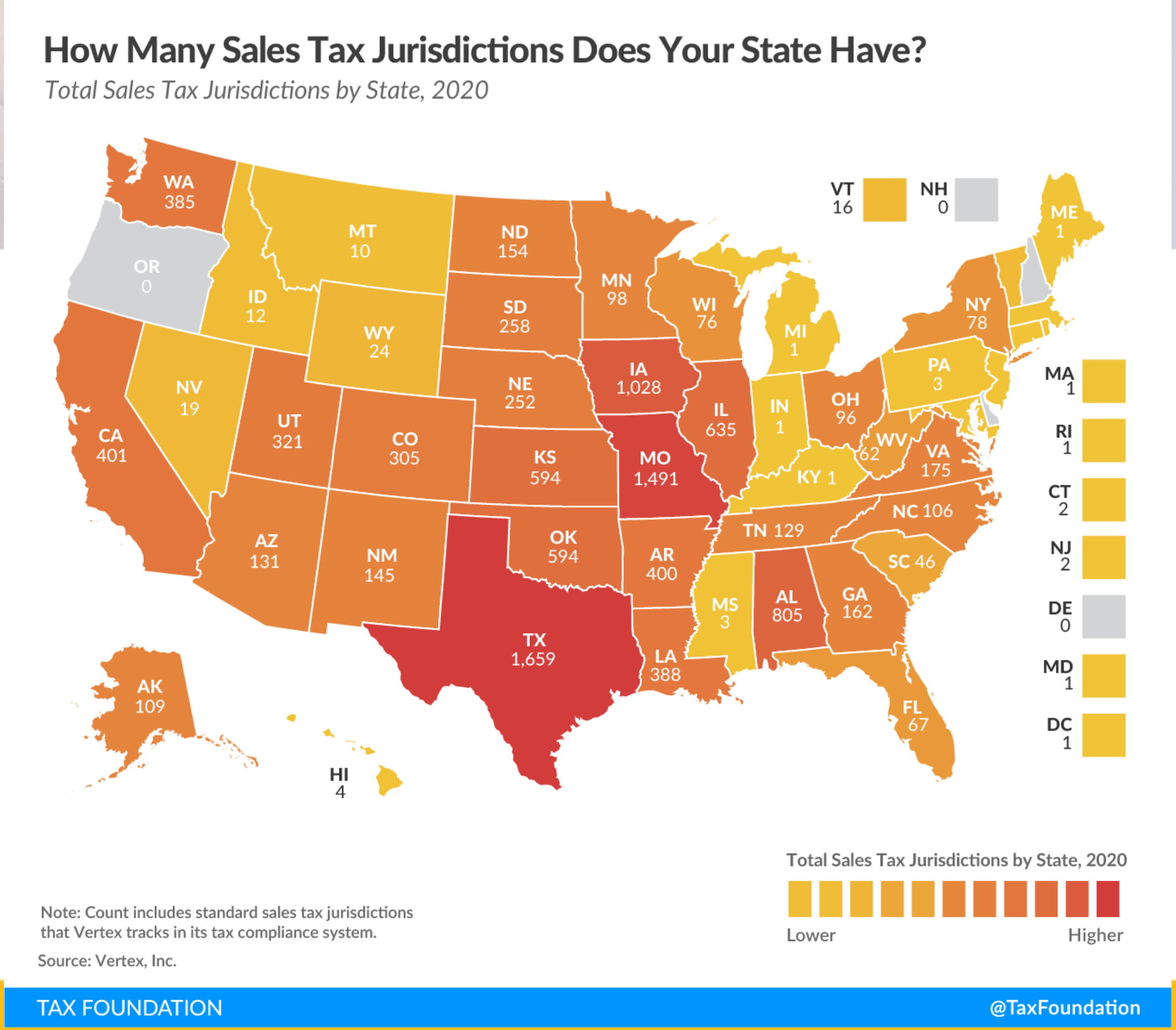
W. P. CAREY SCHOOL of BUSINESS
ARIZONA STATE UNIVERSITY

Research Question

Does local, rather than state, administration of local sales taxes reduce business activity in the real economy?

- **Incremental compliance costs!**

10,000+ *sales* tax jurisdictions across the United States



W. P. CAREY

Who collects these local sales taxes?

- In most states, the *state* department of revenue collects and administers local *and* state sales taxes.
- In some states, however, *local* governments collect and administer the *local* sales taxes.

Table 4						
Locally Administered Sales Taxes – By State and By Type						
State	As of Date	Sales and Use	Rental	Lodging	Food and Beverage	Admissions and Amusement
Alabama	February 2016	327	133	132		
Alaska	January 2016	100		44		
Arizona	February 2016	14				
California	June 2014			465		
Colorado	December 2015	72		49		
Florida	January 2016			40	1	
Georgia	December 2015			267	N/A*	
Idaho	February 2016	7		13	9	
Illinois	June 2015		1	N/A*		N/A*
Indiana	September 2015				1	
Louisiana	February 2016	63				
Maryland	January 2015			26		
Minnesota	October 2015			132		
New York	February 2016			N/A*		
North Carolina	February 2016			159	5	
Ohio	January 2014			410		63
Oklahoma	March 2015			40		
Oregon	February 2016			107		
Pennsylvania	December 2015			60		349
South Carolina	December 2014			122	129	
Tennessee	July 2012/2015			133		3
Vermont	February 2016			2	2	2
Virginia	December 2013			184	190	22
Washington	February 2016					56
West Virginia	February 2016			130		

Incremental Costs

- Filing more returns/payments (manager time/paying accountant/purchasing software)
- Tax information frequently not available from a central source
- Register with each taxing jurisdiction
- Differences between state and local sales tax base
- Enforcement procedures (audits, notices, appeals) dealt with on jurisdiction-by-jurisdiction basis

How Large are Incremental Costs?

- 2016 report prepared by KPMG for the Institute for Professionals in Taxation
  - Aggregate $190 million annual incremental cost for the 5 states they examined (Alabama, Alaska, Arizona, Colorado, and Louisiana)

Prediction: Does local tax administration matter?

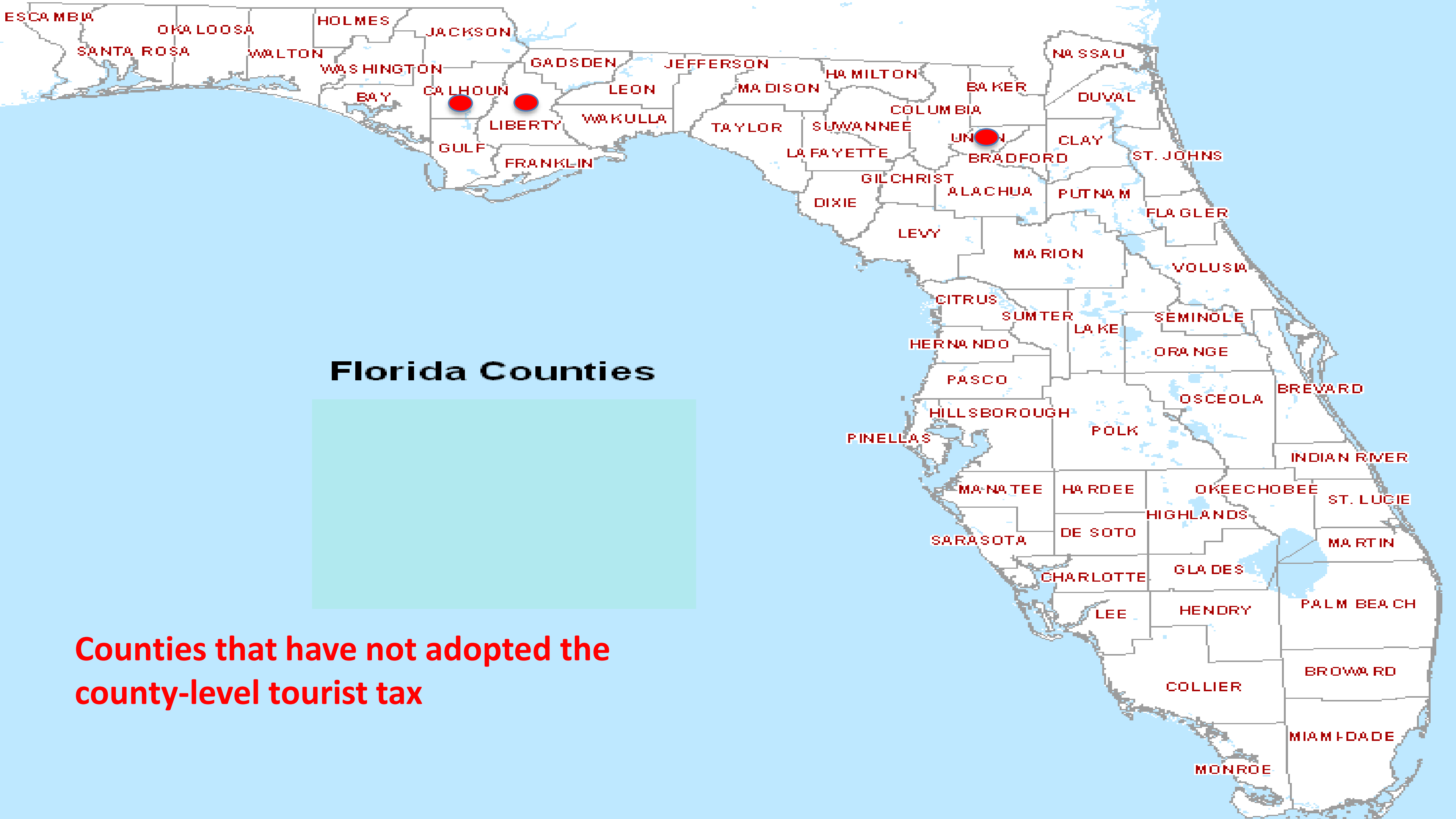
If incremental compliance costs meaningfully cut into earnings, or prospective earnings, then business activity in the real economy will be reduced.

Our Setting

- Florida Tourist Development Tax (tourist or hotel tax)
  - County-level tax on short-term rentals (e.g., hotel rooms, private homes)
  - Originally administered by the state (in conjunction with state tourist sales tax)
  - In late 1980s, counties given choice to continue state administration or county could self-administer

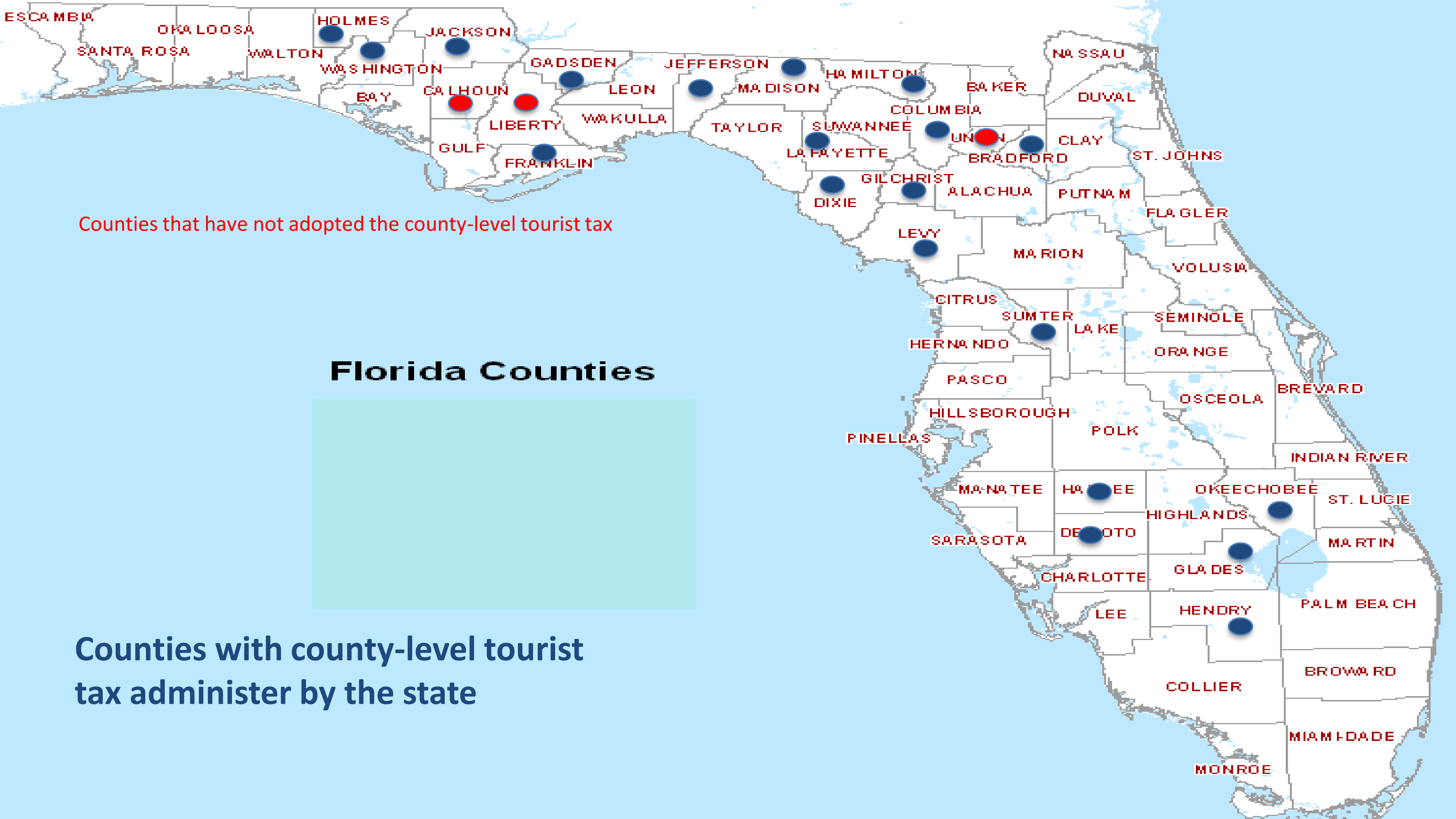
Our Setting

We examine whether switches between state and local sales tax administration have real economic effects on the hotel industry in Florida.



Florida Counties

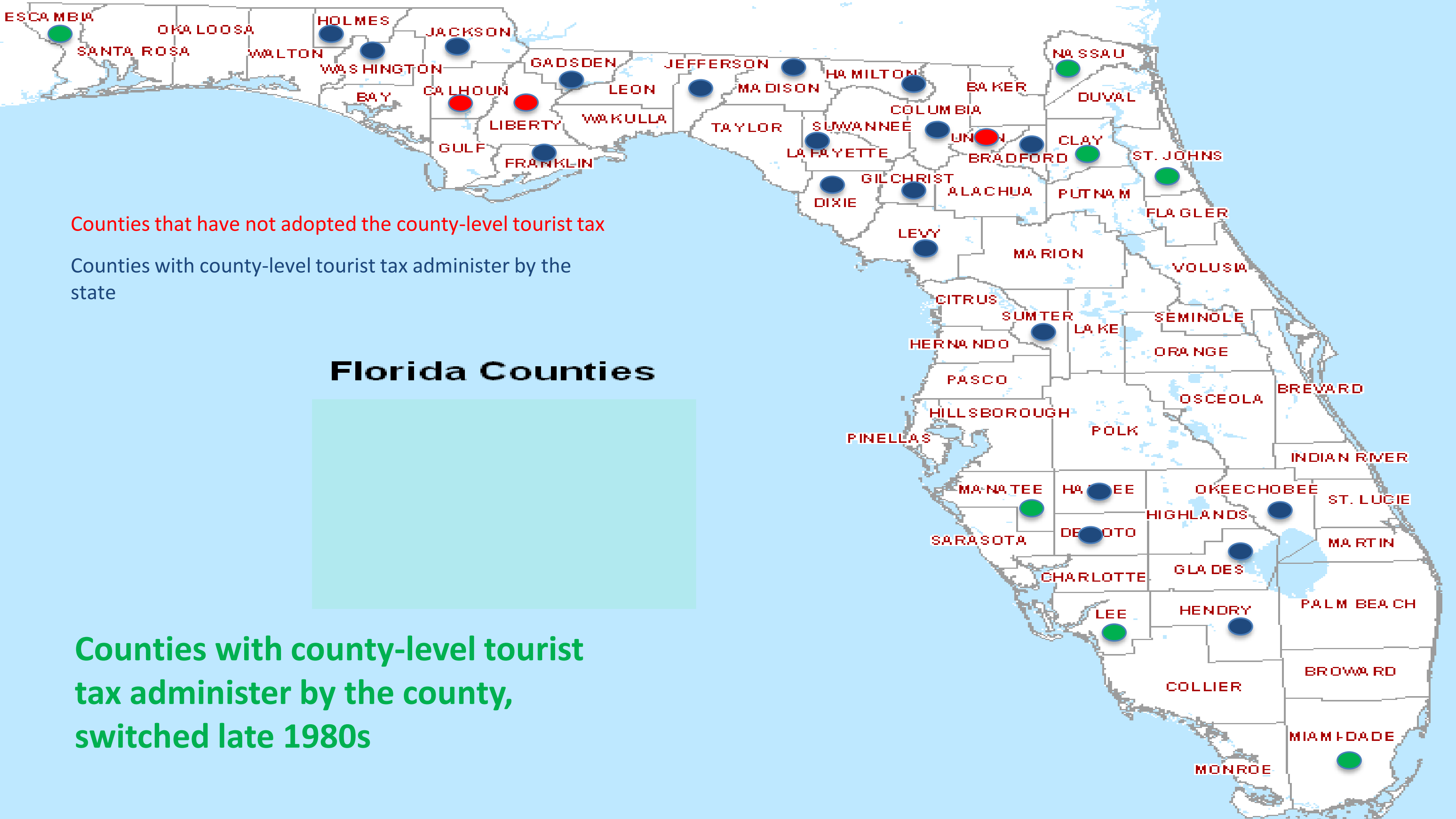
Counties that have not adopted the county-level tourist tax

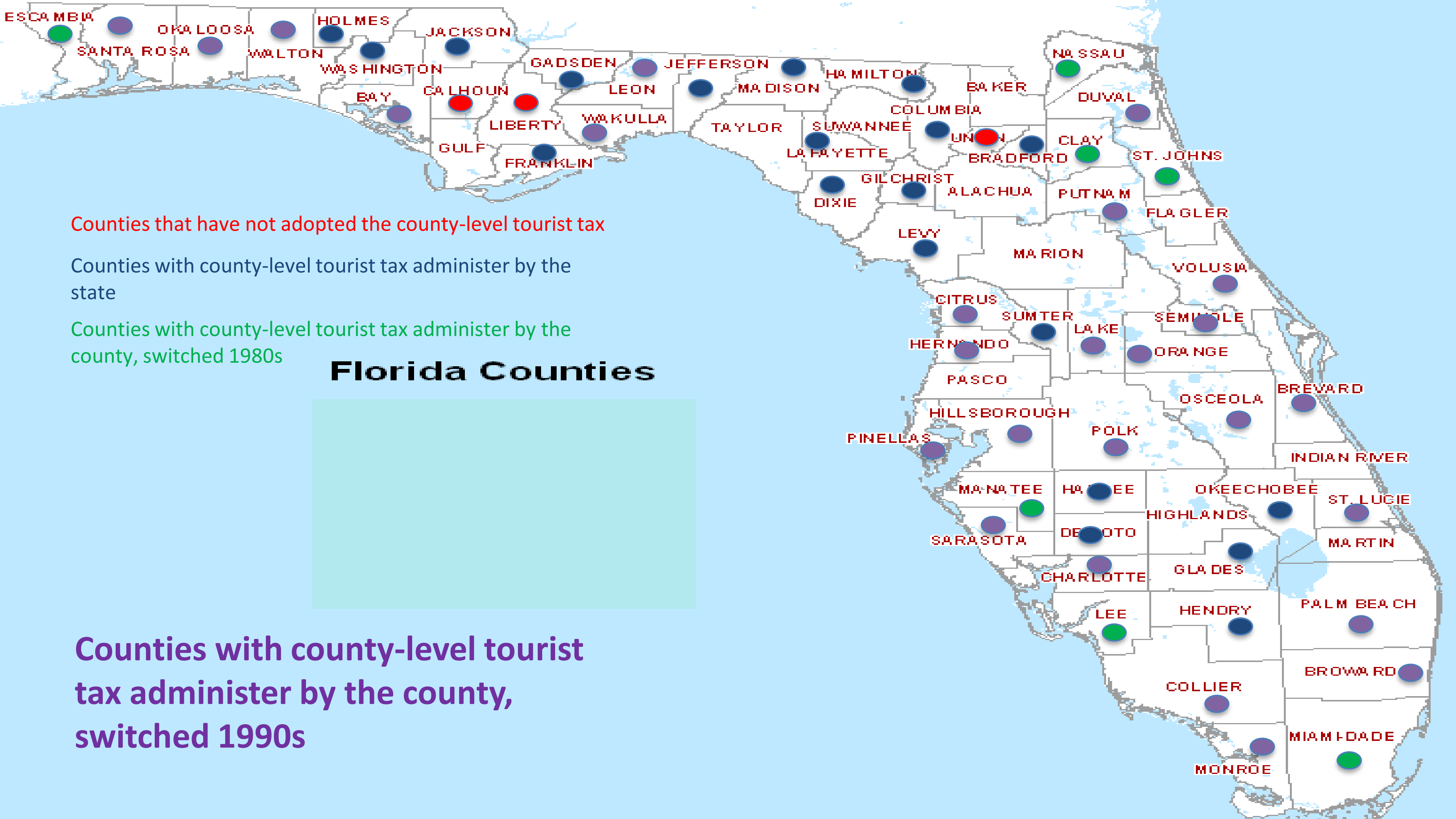


Counties that have not adopted the county-level tourist tax

Florida Counties

Counties with county-level tourist tax administer by the state


Florida Counties
Counties that have not adopted the county-level tourist tax
Counties with county-level tourist tax administer by the state
Counties with county-level tourist tax administer by the county, switched late 1980s
ESCAMBIA
SANTA ROSA
OKALOOSA
WALTON
HOLMES
WASHINGTON
JACKSON
BAY
CALHOUN
GADSDEN
LIBERTY
GULF
FRANKLIN
WAKULLA
LEON
JEFFERSON
MADISON
TAYLOR
HAMILTON
SUWANNEE
LAFAYETTE
COLUMBIA
BAKER
NASSAU
DUVAL
CLAY
BRADFORD
ST. JOHNS
GILCHRIST
ALACHUA
DIXIE
PUTNAM
FLAGLER
LEVY
MARION
VOLUSIA
CITRUS
SUMTER
LAKE
SEMINOLE
HERNANDO
ORANGE
PASCO
OSCEOLA
BREVARD
HILLSBOROUGH
POLK
PINELLAS
INDIAN RIVER
MANATEE
OKEECHOBEE
ST. LUCIE
HIGHLANDS
SARASOTA
MARTIN
GLADES
CHARLOTTE
LEE
HENDRY
PALM BEACH
BROWARD
COLLIER
MIAMI-DADE
MONROE

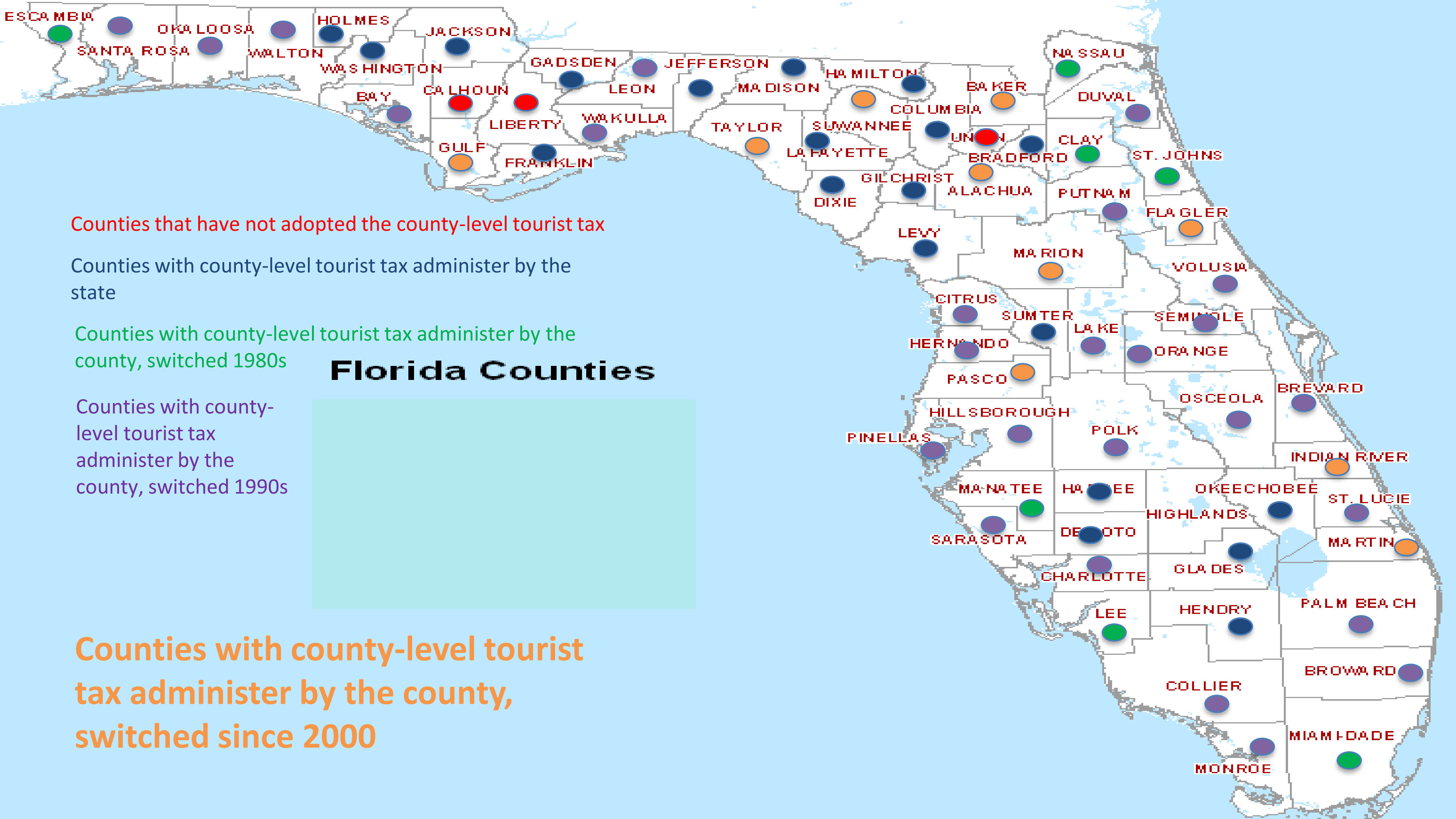


Counties that have not adopted the county-level tourist tax

Counties with county-level tourist tax administer by the state

Counties with county-level tourist tax administer by the county, switched 1980s

Counties with county-level tourist tax administer by the county, switched 1990s



Counties that have not adopted the county-level tourist tax

Counties with county-level tourist tax administer by the state

Counties with county-level tourist tax administer by the county, switched 1980s

Florida Counties

Counties with county-level tourist tax administer by the county, switched 1990s

Counties with county-level tourist tax administer by the county, switched since 2000

Why Switch to Local Administration?

- Receive revenues from the tax more quickly
- Capture jobs related to the collection, enforcement, and auditing of tax
- More control of enforcement
  - Choose whom to audit
  - Improve monitoring of evasion by privately owned rental properties

Identification

- County-level choice to administer county-level tourism tax not endogenously determined by hotel characteristics
    - Such as lack of compliance

- Negative impact of local tourism tax administration on the hotel industry can be attributed to a (plausibly) exogenous, unintended consequence of increased compliance costs.

Research Design

$$Hotel\ Industry\ Size_{c,t} = \alpha + \beta_1 \text{ x } Local\ Hotel\ Tax_{c,t} \text{ x } Locally\ Administered_{c,t} + \beta_2 \text{ x } Local\ Hotel\ Tax_{c,t} + \Sigma\ Controls_{c,t}$$

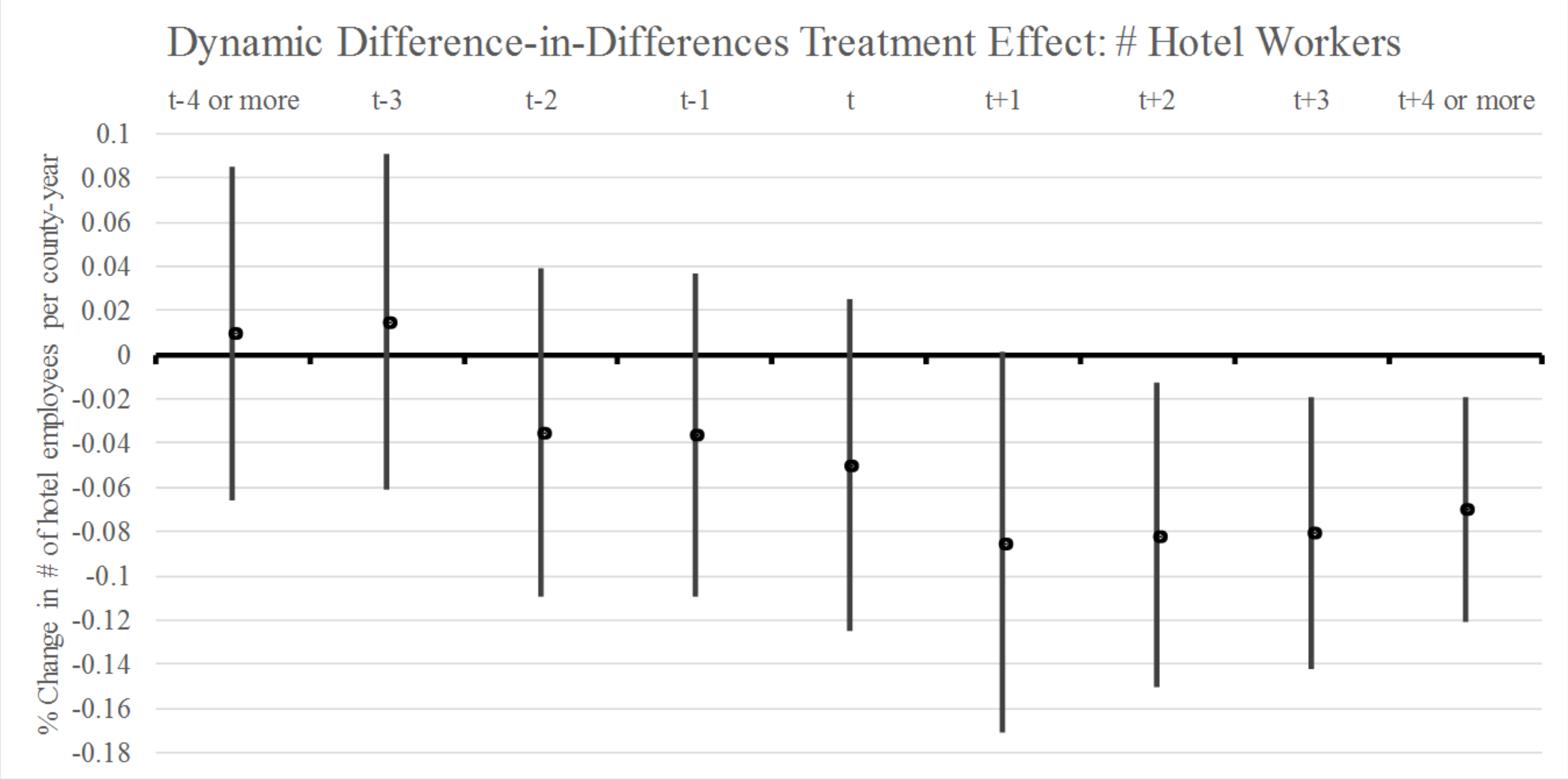
Hotel Industry Size$_{c,t}$: # of hotels, # of hotel workers, or total hotel industry wages in county *c* in year *t*

Data from BLS QCEW. Surveys at county-year-industry level.

1990-2019: 30 year panel of 67 counties for 1,556 county-years

Table 6: Changes in county-year employment

$Ln(1 + \#\ Hotel\ Employees_{c,t}) = \alpha + \beta_1 \times Local\ Hotel\ Tax\ (\%)_{c,t} \times Locally\ Administered_{c,t} + \Sigma\ Controls_{c,t}$	1	2	3
Local Hotel Tax (%)$_{c,t}$ x Locally Administered$_{c,t}$			-0.0528**
			[-2.32]
Local Hotel Tax (%)$_{c,t}$		-0.0529*	-0.0118
		[-1.74]	[-0.34]
Local Sales Tax (%)$_{c,t}$	-0.1052	-0.1138	-0.1175
	[-1.21]	[-1.35]	[-1.42]
Population (1,000s)$_{c,t}$	-0.0001	-0.0001	-0.0001
	[-0.01]	[-0.08]	[-0.01]
# Adjacent County Hotel Employees (1,000s)$_{c,t}$	-0.0117**	-0.0129**	-0.0105*
	[-2.23]	[-2.40]	[-1.90]
# Manufacturing Employees (1,000s)$_{c,t}$	-0.0061	-0.0045	-0.0039
	[-0.99]	[-0.73]	[-0.71]
# Finance Employees (1,000s)$_{c,t}$	0.0001	-0.0004	0.002
	[0.00]	[-0.05]	[0.31]
County and Year Fixed Effects	Yes	Yes	Yes
Observations	1,556	1,556	1,556
R^2	96.18%	96.20%	96.23%


Dynamic Difference-in-Differences Treatment Effect: # Hotel Workers
t-4 or more
t-3
t-2
t-1
t
t+1
t+2
t+3
t+4 or more
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
-0.12
-0.14
-0.16
-0.18
% Change in # of hotel employees per county-year

Conclusion:

Compliance costs of locally administered sales taxes discourage investment, total wages, and employment in the Florida hotel industry

Contributions:

- Sales tax decentralization lowers real economic activity
- Add to accounting literature on the effect of tax administration on the behavior of businesses
- Add to literature suggesting that complex tax systems create deadweight compliance costs that slow economic development

Thank You

Using Discrete Event Simulations to Understand the Impact of Changes to IRS Processes

Deandra Reinhart, Supervisory Social Scientist, IRS
Rafael Dacal, Senior Operations Research Analyst (Presenter), IRS
Ariel S. Wooten, Acting Lead Social Scientist, IRS
Patrick Kaylor, Social Scientist, IRS
Jonathan Curtiss, The MITRE Corporation (IRS Contractor)

June 2021

Background

The Automated Underreporter System (AUR) is an SB/SE program that "systemically identifies potential cases through the computer matching of tax returns with corresponding Information Returns Master File (IRMF) taxpayer information documents" (Internal Revenue Manual (IRM) 4.19.3.1.1).

- The identified or matched cases are further reviewed by a Tax Examiner (TE) to evaluate for underreporting and/or over-deduction discrepancies.

In the past, to test an organizational process change, SB/SE disrupted the existing system to conduct a pilot test.

- Pilots cause down time and lost productivity
- Discrete Event Simulation (DES) allows SB/SE Research to predict the impact of potential changes to different operational aspects without incurring program down time.

This project is a proof of concept, to evaluate if DES might be beneficial for other SB/SE work units, in addition to AUR.

Objective, Market Segment, and Sample Frame

AUR DES Research Objectives:

- Assess the efficacy of DES in the context of the AUR operation
- Estimate how changes in AUR resources or processes might impact AUR operations under experimental scenarios.

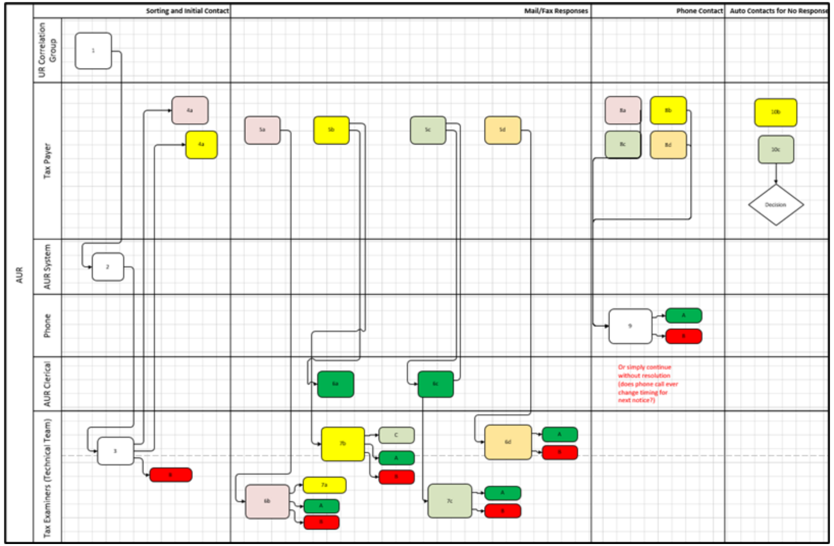
Market Segment and Data Sample

- AUR clerical employees and TEs at the Philadelphia Campus served as Subject Matter Experts (SMEs)
- Taxpayers in the Philadelphia Campus AUR inventory for tax years 2016-2019 are included in this DES.

Business Process Model (BPM)

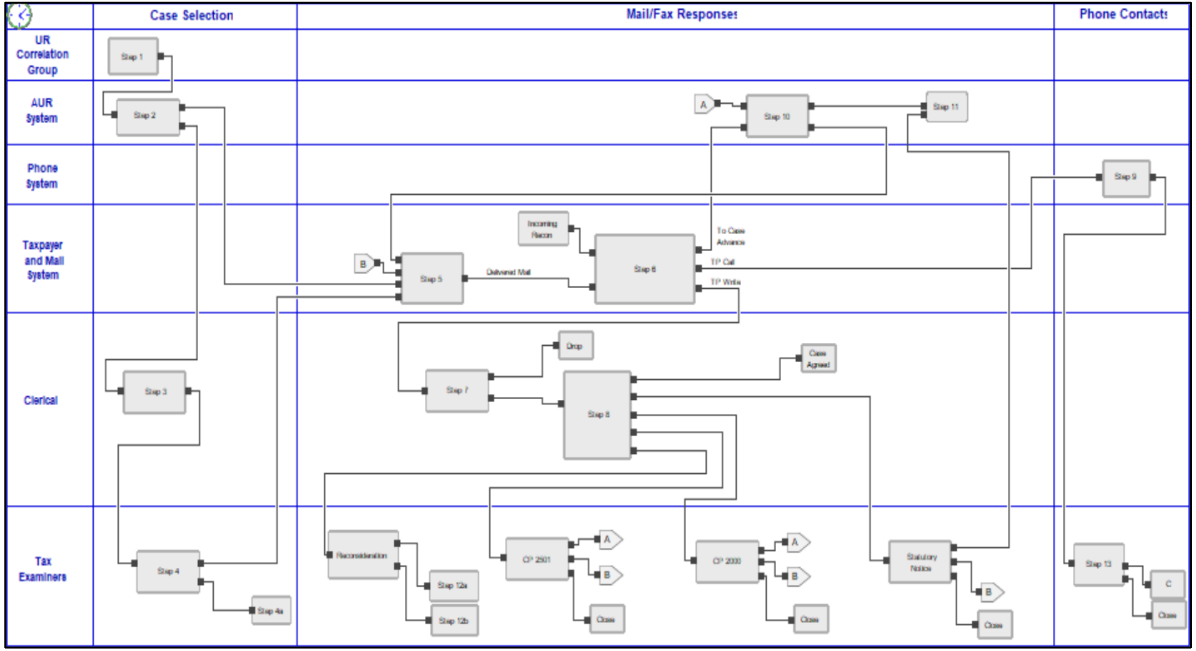
- How cases/modules move between interconnected processes.
  - Internal Revenue Manual as resource
  - Subject Matter Experts (SMEs) on team
- Business Process Modeling and Notation (BPMN) standards
  - Consider and document all the potential paths by which a case may travel through the process (advance or fail to advance included as well)
- Evaluate the configuration of all inputs, path decisions, timing, resources, and outputs in the model
- Activities are contained in “swim lanes” to indicate responsibility for independent business processes.

AUR Process Flow



Executable BPM Creation

After BPM completion and development of a full process flow, move the model to simulation software.



Fidelity and Verification

Measuring Fidelity

- Fidelity is “the degree to which a model or simulation reproduces the state and behavior of a real-world object or the perception of a real-world object, feature, condition, or chosen standard in a measurable or perceivable manner; a measure of the realism of a model or simulation; faithfulness.”

Completing Model Verification

- Ensure the model is constructed correctly
- Verify that the conceptual model accurately represents the real-world data
- Begin with a simple model and add complexity as needed.
  - Ensure group involvement among developers to check one another’s work.
  - Conduct calculation checks.
  - Run simple model, where the assumptions or conditions are known to be true, with added complexity

Validation and Calibration

Validation

- Examine Model Accuracy and Data Representativeness
  - Users of the model determine acceptable level of accuracy
  - Results should permit end-users to make informed decisions
- Two activities for validating a model
  - Validate conceptual model (often a qualitative process)
  - Test the model’s accuracy (an iterative processes, with SME input)

Calibration

- Compare the model to the system being simulated
- Adjust incorrect model parameters to improve calibration
  - Done in conjunction with the validation process

Inputs and Outputs

Inputs

- Real world values and metrics of interest obtained from the live AUR operation
  - Inserted into the model in segmented phases.
    - Phases indicate where information input is necessary for the model to continue to the next step in the AUR process.
- Type of inputs:
  - AUR inventory metrics such as taxpayer response rate(s), volumes of mailed notices, inventory age
  - IRS personnel are arranged into pools (groups of resources available to work specific parts of a process)
  - Time (such as shifts, estimated time for activity completion or taxpayer response)

Output Logs

- The executable model's estimates at the beginning and the end of identified activities.
- Calculate changes in AUR inventory metrics to validate and calibrate the model

Experiments

SB/SE Research conducted two experiments to validate the AUR DES model.

1. Post-COVID restart experiment
   - Assess bottlenecks in the AUR process created by the sudden campus closures.
   - Calculate the time required to clear and process the correspondence backlog given different restart assumptions.
   - Forecast taxpayers’ responses and downstream impact on tax examiners’ case workload.
2. Notice redesign experiment - evaluate the impact of a potential 10, 20, and 30 percent increase in taxpayer notice response rates.

Findings and Benefits of Simulation

This proof of concept demonstrates the potential effectiveness of DES to illustrate workflows, and test operational changes without workflow disruption within the IRS environment.

We successfully:

- Provided insights to identify workflow bottlenecks
- Projected the impact of changes in work volumes and staffing
- Estimated the optimal use of resources to meet work demands without impacting existing work processes.
- Suggested future research surrounding process changes
  - Estimate how changes to the automated routing calling system may impact the overall process
  - Estimate how changes to the batching process may impact the process
  - Test impact of notice redesigns
  - Estimate how resource changes (increase or decrease) within organizations or individual campuses impact inventory

Conclusion

1. **The AUR Philadelphia simulation model and subsequent experiments demonstrated that DES can be used to estimate the impact of changes to complex processes while minimizing disruptions to existing systems.**

2. **Post COVID-19 campus restart experiment results enabled AUR leadership to identify potential risks and develop mitigation strategies.**

3. **Model development and experiments showed that DES can be useful throughout the tax administration community.**

Annual IRS-TPC Joint Research Conference on Tax Administration

June 24, 2021

Effects of Forms 1040-X and 1120-X Post-filing Adjustments on SOI Estimates

by Derrick Dennis, Jennifer Ferris, Gloria “Chloe” Gagin, Tuba Ozer-Gurbuz, Julia Shiller, and Christopher Williams

Introduction

- Historically, Statistics of Income (SOI) research has been based on data collected from initial tax return filings
  - Although initial tax return filings are often final, taxpayers may elect to amend their income tax returns
    - Common reasons for amending an initial return include tax over- or underassessments, or the taxpayer might file an amendment after identifying errors on the original return

Introduction

- Taxpayers have an option to amend their initial return after the tax filing period has ended
- Under Internal Revenue Code (IRC) § 6511, taxpayers must file a claim for a credit or refund within 3 years from the time the initial tax return was filed or 2 years from the time the tax was paid, whichever is later
- In some instances, extensions may be granted for filing an amended return

Scope of Project

- Estimate the effect of including amended tax returns into SOI products

  - *U.S. Individual Income Tax Return* (Form 1040)
  - *U.S. Corporation Income Tax Return* (Form 1120)

  - *Amended U.S. Individual Income Tax Return* (Form 1040-X)
  - *Amended U.S. Corporation Income Tax Return* (Form 1120-X)

Scope of Project

- Base year SOI Year 2013
  - This means that this study measures changes made in subsequent years to returns that were filed for SOI Year 2013
    - a relatively recent year with 3 years of post-filing data

- Tax adjustments were identified in IRS Master File data and linked to tax returns in the individual and corporate SOI samples
    - The matched data were weighted using the respective sample weights from the individual and corporate studies to represent adjustments in the overall population
    - Weighted sample estimates were compared to total changes in the population data to verify the appropriateness of SOI sample weights for this study

Results

- Individual Income Tax Returns

Table 1. Effects of Post-filing Adjustments on Forms 1040, SOI Year 2013
[Money amounts are in millions of dollars]

Item	Amendments [1]
	Recommended tax change
Increases to tax	$2,500
Decreases to tax	$4,600
Net impact on SOI estimates [2]	**$-2,100**
Percent effect on SOI estimates	**-0.16%**

[1] Amendments refers to post-filing adjustments induced by a Form 1040-X.
[2] This row is calculated from approximately 3.2 million amended returns.
SOURCE: IRS, Statistics of Income Division, Weighted Sample, February 2021.

Results

Figure B

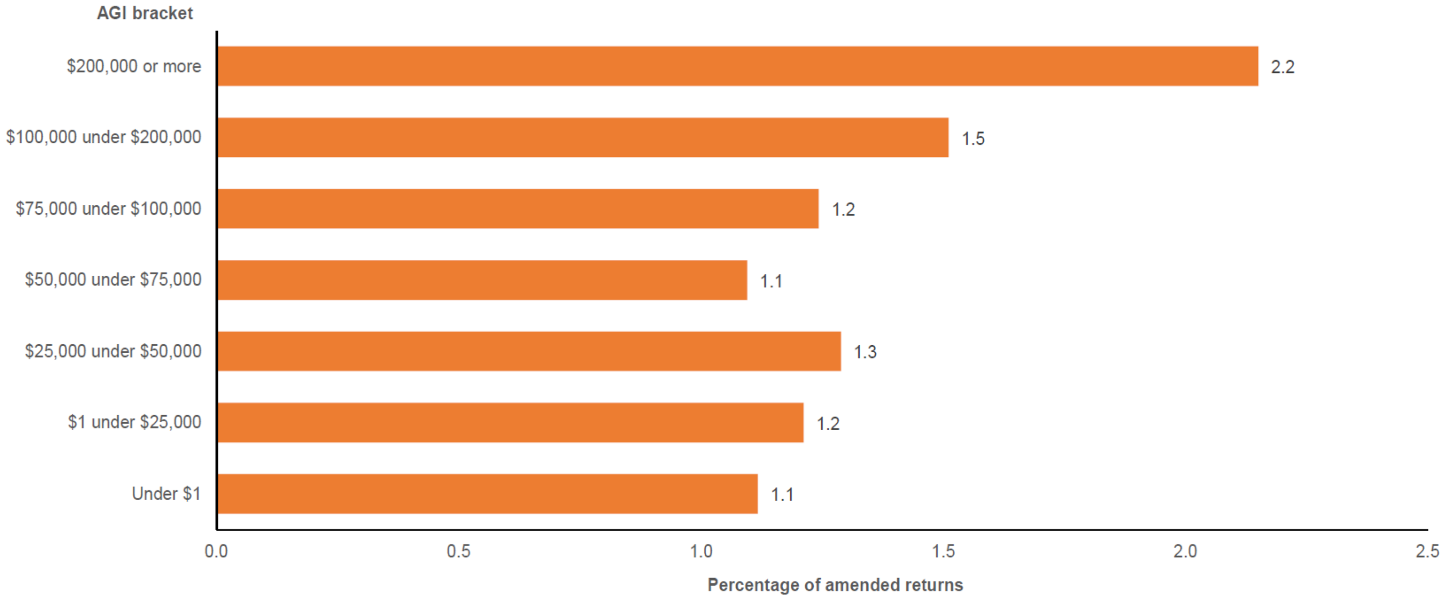
Form 1040-X Filing Rates by Filing Status, Tax Year 2013

SOURCE: IRS, Statistics of Income, Weighted Sample, February 2021.

Results

Figure H

Form 1040-X Filing Rates, Married Filing Jointly Filers, by Selected Adjusted Gross Income Bracket, Tax Year 2013



NOTE: Percentage labels were rounded.
SOURCE: IRS, Statistics of Income Division, Weighted Sample, February 2021.

Results

- Corporate Income Tax Returns

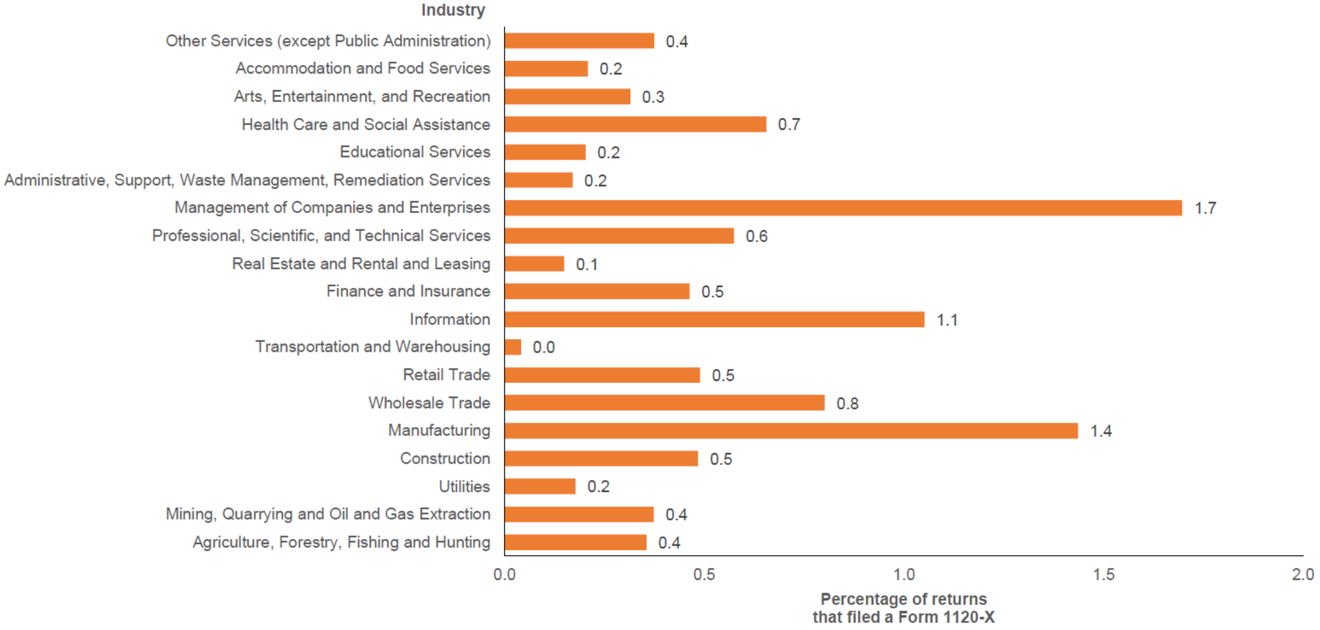
Table 2. Effects of Post-filing Adjustments on Forms 1120, SOI Year 2013
[Money amounts are in thousands of dollars]

Item	Amendments [1]	
	Number of returns	Recommended tax change
Increases to tax	1,251	$30,397
Decreases to tax	1,778	$118,200
Net impact on SOI estimates [2]	**3,029**	**$-87,803**
Percent effect on SOI estimates	**0.16%**	**-0.03%**

[1] Strictly includes returns intially filed with Form 1120.
[2] Returns with no change to tax are not included.
SOURCE: IRS, Statistics of Income Division, Weighted Sample, February 2021.

Results

Figure I
Form 1120-X Filing Rates by Industry, SOI Year 2013

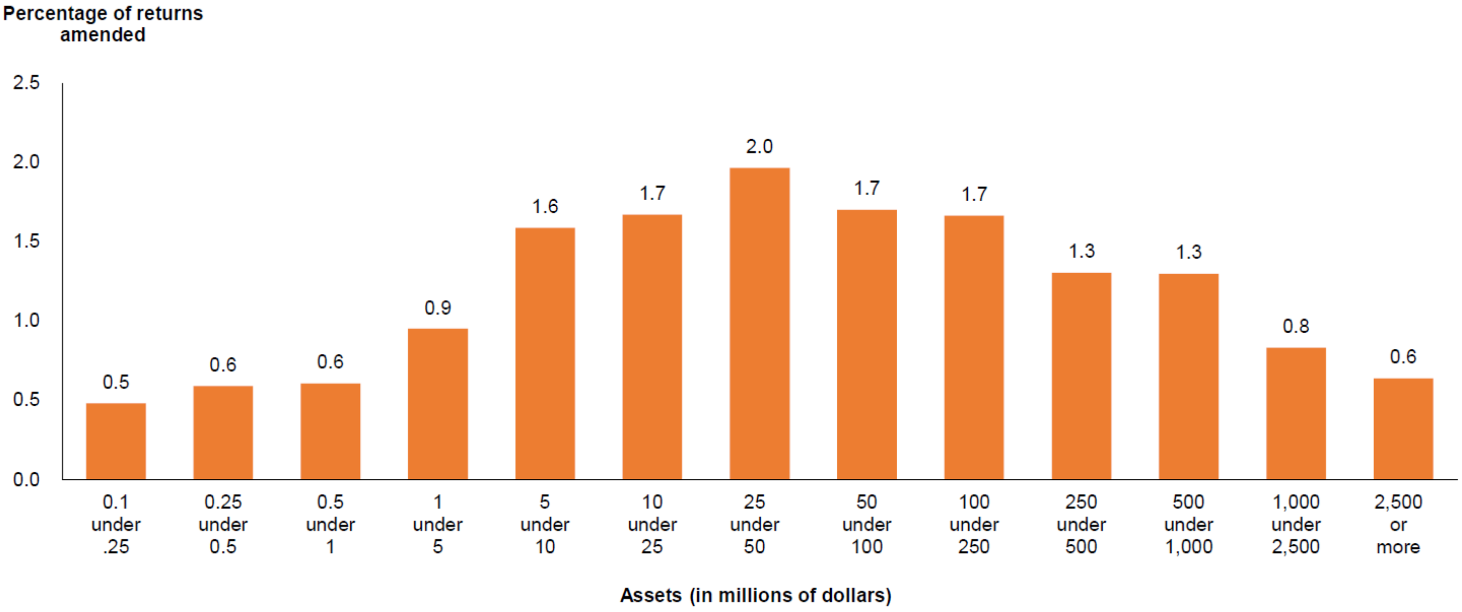


NOTE: Percentage labels were rounded.
SOURCE: IRS, Statistics of Income Division, Weighted Sample, February 2021.

Results

Figure J

Form 1120-X Filing Rates by Size of Total Assets, SOI Year 2013



NOTE; Percentage labels were rounded.
SOURCE: IRS, Statistics of Income Division, Weighted Sample, February 2021.

Conclusion

- In this preliminary study looking at 2013,
  - Some 2.3 percent of individual income tax returns (Forms 1040) and 0.52 percent of corporate income tax returns (Forms 1120) were amended (using Forms 1040-X and 1120-X)
  - These amendments decreased the estimated total tax from individual filers by only $2.1 billion (0.16 percent) and decreased the estimated total tax from corporate filers even less, by $87.8 million (0.03 percent).
- Thus, the study's review of individual and corporate tax returns from 2013 suggests that while SOI tax return statistics would be affected by post-filing adjustments, current SOI statistics do not significantly change.

Thank You

Gloria “Chloe” Gagin
Gloria.M.Gagin@IRS.gov

Impacts of Variations in Process Discussant Comments

Jennifer Stratton*

*The opinions expressed in this presentation are the author's own and do not necessarily reflect the views of the US Government Accountability Office.

Overview: administration matters

- Papers highlight the importance of tax administration
  - Its potential effect on the economy, beyond policy
  - How to better improve it
  - How challenges in administration effect available information.

Sales Tax Administration and the Real Economy
Potential Correlations

- Correlation with Hotel Industry Size and choice to administer taxes locally?
  - Could localities with smaller sized industries choose to administer locally because they don't expect to have a lot of challenges?
- Correlation with Hotel Industry Size and share of smaller transient renting venues (Airbnb)?
  - As noted in the paper, smaller taxpayers—individuals renting lone vacation homes for short-term stays—can evade state-administered tourist taxes. County officials are both more informed about these smaller operators and have stronger incentives to enforce tax compliance
  - Suggests a possible omitted variable, that those counties that have a lot of small taxpayers renting rooms are both more likely to locally administer and have smaller hotel industry size. This has also been an increasing trend that may dominate the decision to locally administer and be negatively correlated with the size the hotel industry.

Sales Tax Administration and the Real Economy
Effective Controls

- Is adjacent county control an effective control for regional hotel industry trends?
  - One of the largest tourist counties is bounded by both another likewise tourist county and one that is nothing but the everglades.
  - More detail on how this effectively controls for a new tourist attraction, which often shows up in already tourist driven areas, or a hurricane which could effect only one county—e.g. Miami-Dade.
  - Both of these types of industry drivers also appear particularly time sensitive. An announcement of a new attraction would likely be reflected over the next few years. The effect of a hurricane also reflected in following years, not necessarily the current year.

Sales Tax Administration and the Real Economy
Alternate conclusion

- The authors conclude that the negative affect of local administered taxes on payroll and employees (but weekly on size) indicate economic affects of administrative costs in the reduction of the size of the hotel industry.
- However, this finding could be the result of a switch from in-house (payroll and employees) tax administration work to contracting with an accounting firm (not payroll or employees, but expenses) to deal with the new multiple filings and deadlines.
  - This type of switch would also be consistent with little findings on establishments.

Using Discrete Event Simulations to Understand the Impact of Changes to IRS Processes

- Quibbles :
  - Limitation to Philadelphia—a reason to think the verification and validation of the model would vary by campus location?
  - Is COVID-19 too big of a change to use as one of the main testing scenarios
- Questions
  - The model includes mail volume and taxpayer response rates, but are any outside factors included, such standard mail times vs. disrupted mail times?
  - In the notice redesign there is an implicit assumption that the redesign increased response rates, how would different assumptions be integrated?

Effects of Post-filing Adjustments on Statistics of Income (SOI) Estimates
Using 2014 Amended returns

- One year limit
  - What is the likelihood that only one year doesn't capture all the amended returns related to a given return?
- Following year limit
  - Concerned that by only looking at the following calendar year, there could be more amended returns filed in the second calendar year.
  - Taxpayers may be triggered to amend their 2013 return *after* completing their 2014 return and noticing something they did wrong. That would be an amended return in CY 2015.

Effects of Post-filing Adjustments on Statistics of Income (SOI) Estimates

Findings

- Qualify conclusions
  - The authors conclude about differences in the likelihood of filing amended returns by different categories, need to know if these are statistically different.
- Other estimates
  - Key return items that most affected SOI statistics
  - Most common amendments

Session 2. Impacts of Variations in Process

Moderator:	***Mary-Helen Risler*** *IRS: RAAS*
Sales Tax Administration and the Real Economy	***Roger White*** *Arizona State University*
Using Discrete Event Simulations to Understand the Impact of Changes to IRS Processes	***Rafael Dacal*** *IRS: SB/SE*
Effects of Post-filing Adjustments on Statistics of Income (SOI) Estimates	***Chloe Gagin*** *IRS: RAAS*
Discussant:	***Jennifer Stratton*** *Government Accountability Office*

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

11th Annual IRS/TPC Joint Research Conference on Tax Administration

Keynote address begins at 12:40 EDT

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

11th Annual IRS-TPC Joint Research Conference on Tax Administration

Keynote Speaker

Mark Mazur

Acting Assistant Secretary for Tax Policy,
US Department of the Treasury

Session 3. Developments in Technology and Analytics

Moderator:	***Terry Ashley*** *IRS: Taxpayer Advocate Service*
New Approaches to Estimating the Extent of Nonfiling	***Alan Plumley*** *IRS: RAAS*
Using Uplift Modeling to Improve ACS Case Selection and Compliance Outcomes	***Jan Millard*** *IRS: RAAS*
Recent IRS Discriminant Function (DIF) Model Improvements	***Getaneh Yismaw*** *IRS: RAAS*
Discussant:	***Brian Erard*** *Brian Erard & Associates*

June 24, 2021

New Approaches to Estimating the Extent of Individual Income Tax Nonfiling

IRS-TPC Research Conference

Alan Plumley (IRS)

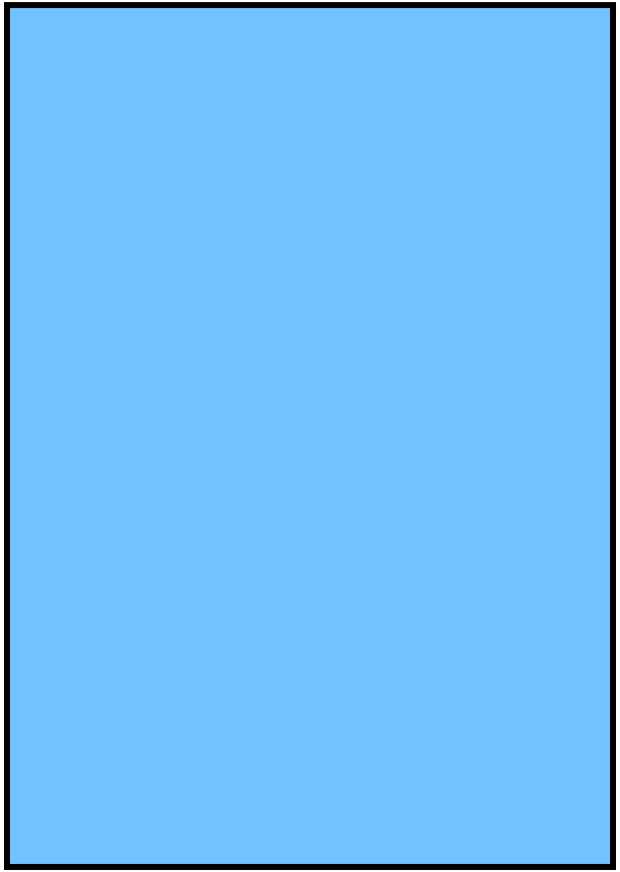
With Tom Hertz, Pat Langetieg, and Mark Payne (IRS), and Maggie Jones (U.S. Census Bureau)

Defining the Extent of Nonfiling

Individual income tax nonfiling tax gap:

The amount of tax not paid on time by those who do not file a required tax return on time.

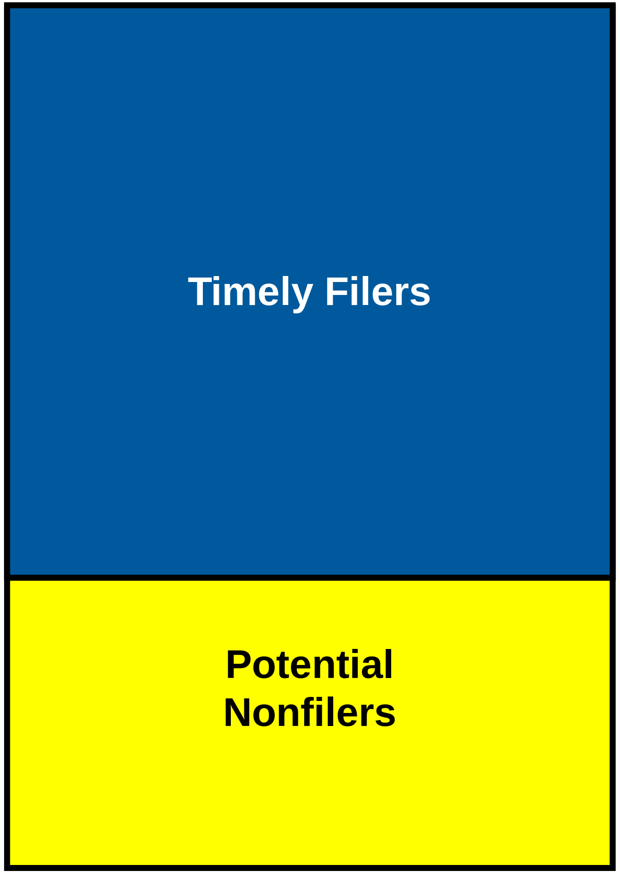
- It is net of any tax that was paid on time by nonfilers (e.g., through withholding, estimated payments, etc.).
- Some nonfilers do not contribute to the nonfiling tax gap.
  - Required to file, but do not have a tax liability or
  - Have paid it in full on time.



Definitions

IRS population **241 M**

Those on a tax form for **Tax Year 2010**



Definitions

IRS population **241 M**

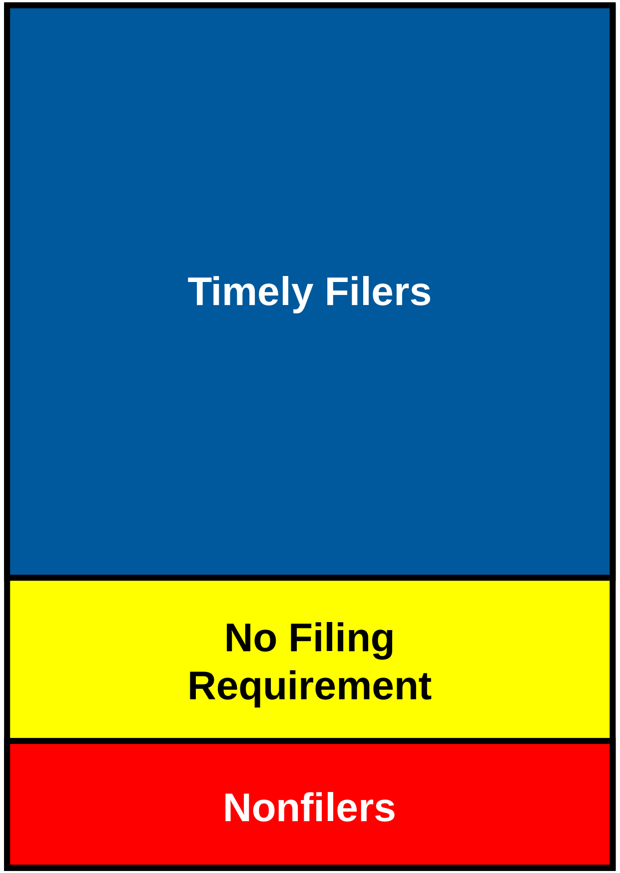
Those on a tax form for **Tax Year 2010**

Timely filers **– <u>190 M</u>**

A primary or secondary on a tax return

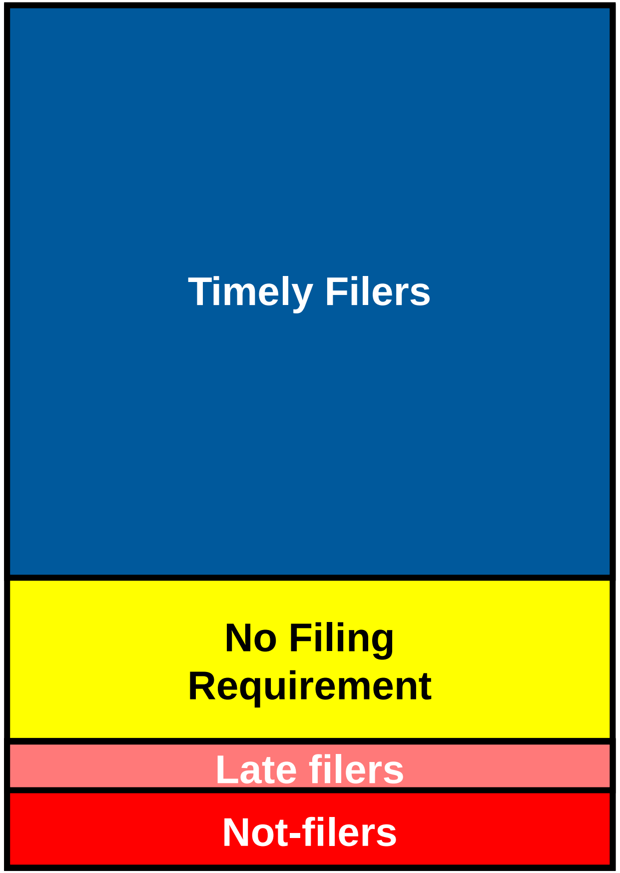
Potential nonfilers **51 M**

People with an information document



Definitions

IRS population	**241 M**
Those on a tax form for **Tax Year 2010**	
Timely filers	**– <u>190 M</u>**
A primary or secondary on a tax return	
Potential nonfilers	**51 M**
People with an information document	
No filing requirement	**– <u>32 M</u>**
Income < thresholds	
Nonfiling individuals	**19 M**
Have a filing requirement	
Nonfiling returns	**15 M**
Married couples combined	



Definitions

IRS population	**241 M**
Those on a tax form for **Tax Year 2010**	
Timely filers	**– <u>190 M</u>**
A primary or secondary on a tax return	
Potential nonfilers	**51 M**
People with an information document	
No filing requirement	**– <u>32 M</u>**
Income < thresholds	
Nonfiling individuals	**19 M**
Have a filing requirement	
Nonfiling returns	**15 M**
Married couples combined	
Late filers	**4 M**
Not-filers	**11 M**

Two Filing Thresholds

Filing Threshold
Gross Income Threshold • Standard deduction + aged/blind deductions + personal exemptions • Depends on **filing status** (single, married-joint, head of household, etc.) • **Problem:** IRS doesn't know filing status of potential nonfilers.

Two Filing Thresholds

Filing Threshold
Gross Income Threshold • Standard deduction + aged/blind deductions + personal exemptions • Depends on **filing status** (single, married-joint, head of household, etc.) • **Problem:** IRS doesn't know filing status of potential nonfilers.
Net Self-Employment Earnings > $433 • **Problem:** IRS generally has little or no information about the self-employment income of potential nonfilers.

Two Filing Thresholds

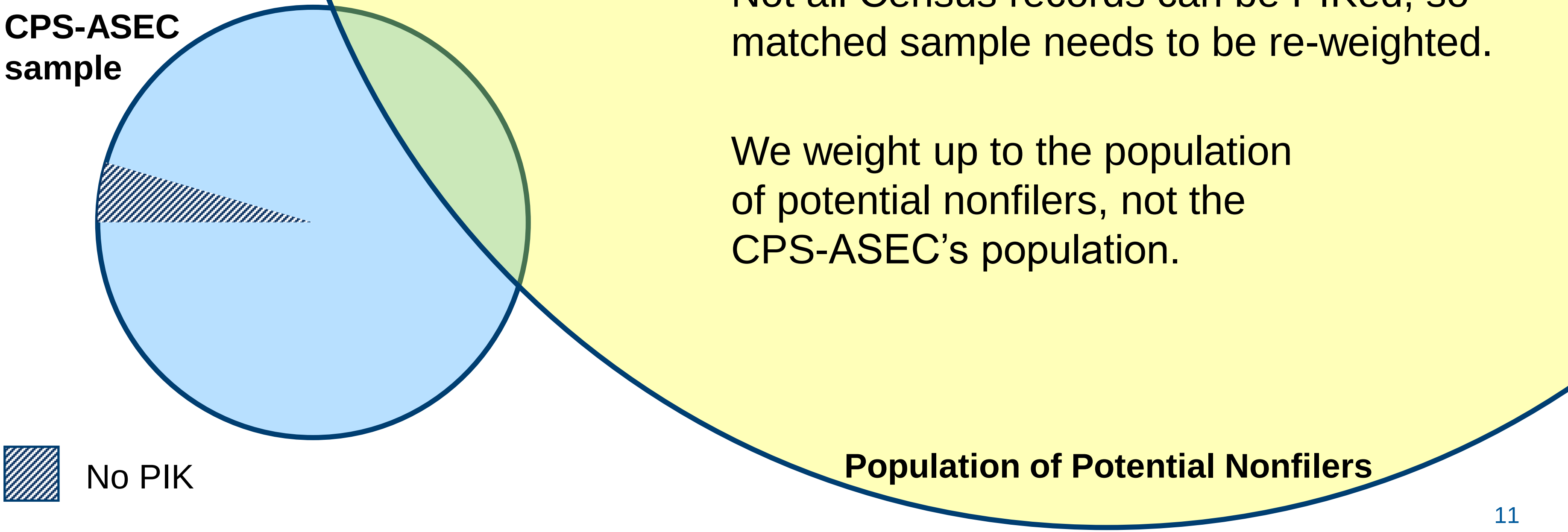
Filing Threshold	Administrative Method (Population Data)
Gross Income Threshold • Standard deduction + aged/blind deductions + personal exemptions • Depends on **filing status** (single, married-joint, head of household, etc.) • **Problem:** IRS doesn't know filing status of potential nonfilers.	• Impute filing status (and dependents) to potential nonfilers using aggregate Census estimates and distribution among filed returns • Sum the income amounts reported by 3rd parties for potential nonfilers • Impute net SE income to potential nonfilers based on net SE income reported on filed returns • 3rd-party income + SE income = total income • Apply both filing thresholds to identify nonfilers **Question:** how reliable is the filing status imputation?
Net Self-Employment Earnings > $433 • **Problem:** IRS generally has little or no information about the self-employment income of potential nonfilers.	

Census Sample Matched to IRS Administrative Data

Comprehensive IRS data made available at Census for tax research. Census Bureau assigns a Protected ID Key (PIK—a unique identifier) to all IRS and Census data for anonymous matching.

Not all Census records can be PIKed, so matched sample needs to be re-weighted.

We weight up to the population of potential nonfilers, not the CPS-ASEC's population.



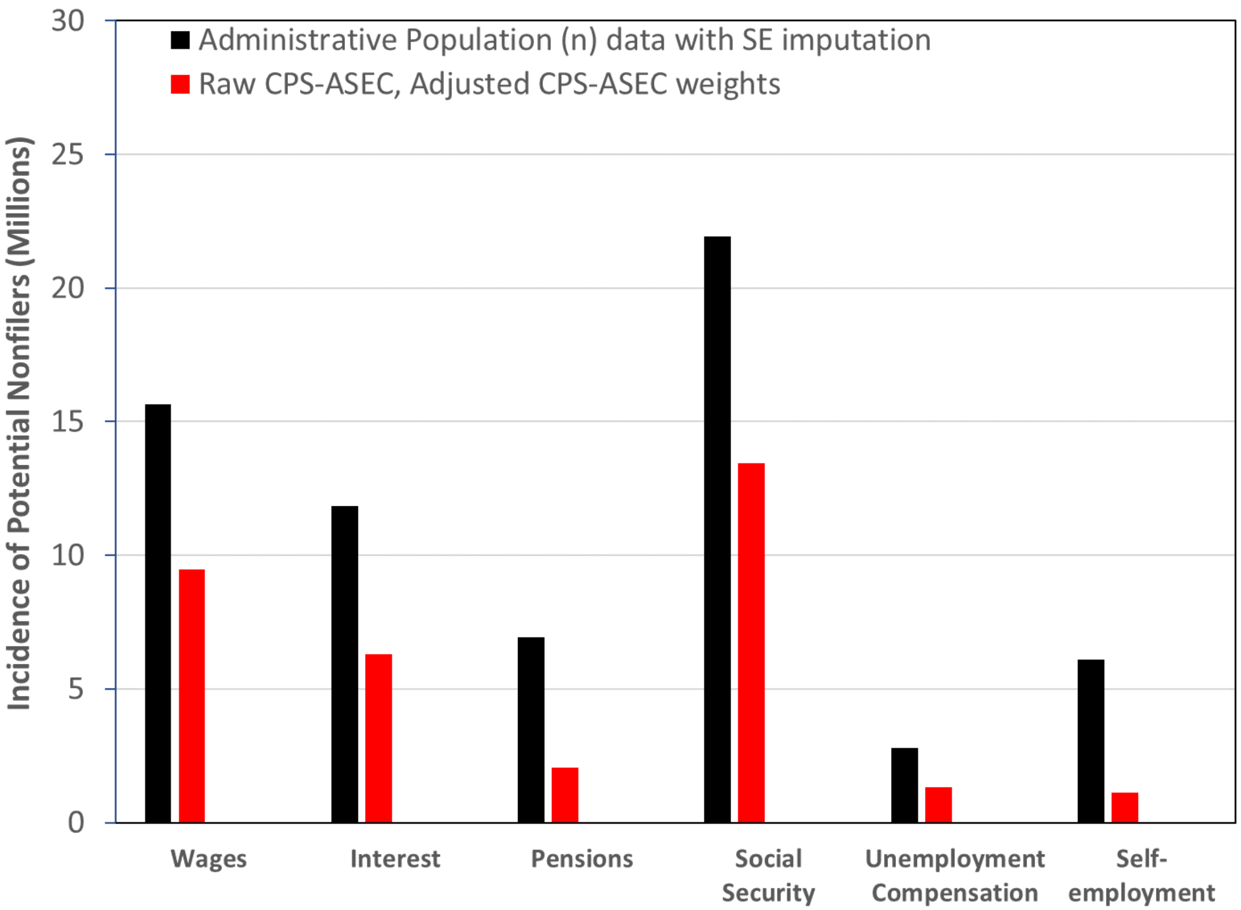
Matching IRS and Census Data: Best of Both Worlds?

Use CPS-ASEC for	Use IRS Administrative data for
• Filing status and dependents at the micro level • Income of spouses who could not be PIKed	• All income reported by 3rd parties • Imputation of net self-employment income • Derivation of new weights to represent population of IRS potential nonfilers

Two main innovations in this paper

1. Improved imputation of self-employment income
   - Controlling better for age of taxpayer
   - Training model on per-exam data from NRP
2. Inclusion of late filers to arrive at comprehensive estimate of nonfiling tax gap

Using Census data and weights under-counts nonfiler income

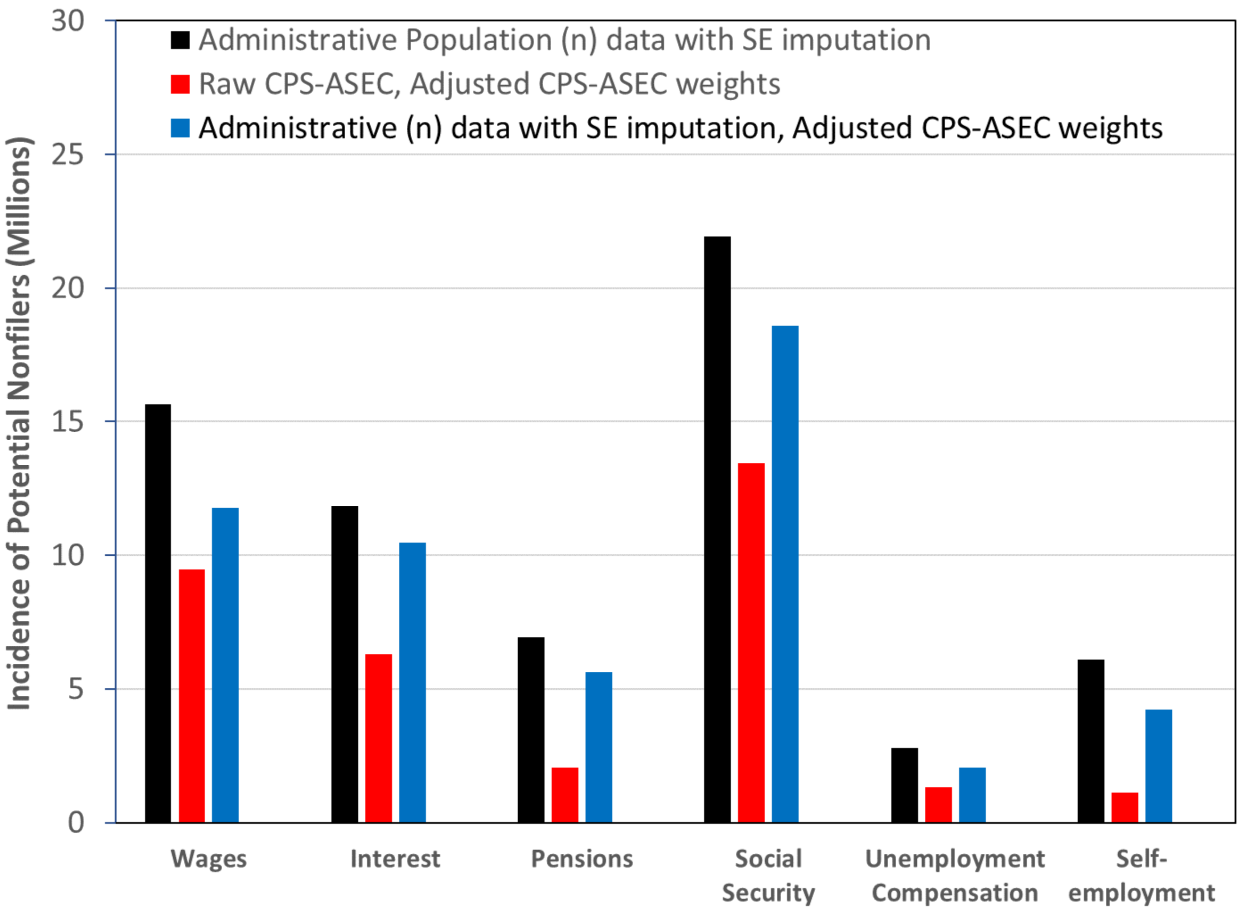


Not all individuals in the CPS-ASEC sample could be PIKed, so the matched sub-sample has to be re-weighted.

Typical adjusted CPS-ASEC weights weight up to CPS-ASEC population.

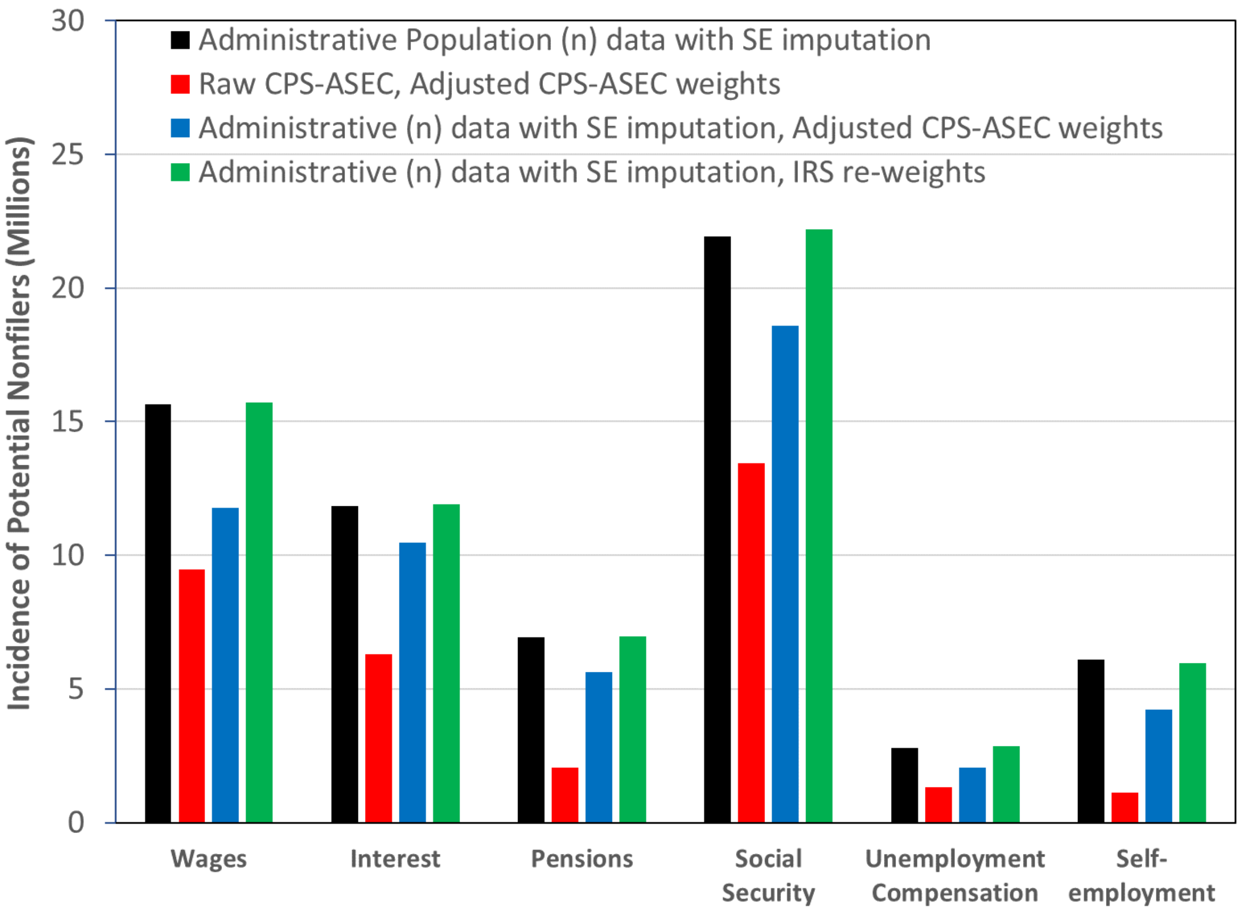
That method applied to CPS-ASEC alone under-counts potential nonfilers with key income types.

Using IRS data and Census weights does better



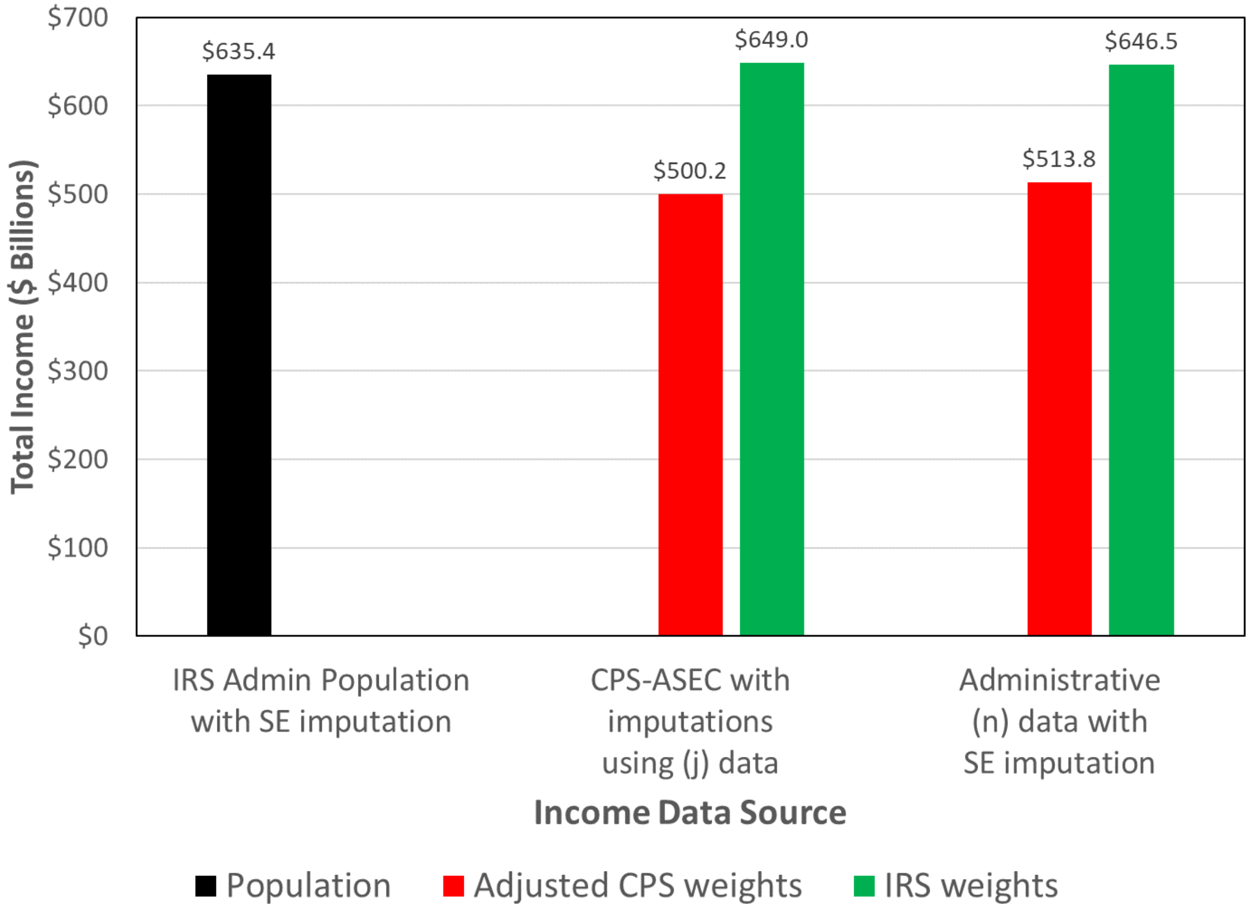
Relying instead on comprehensive IRS administrative data (with the SE imputation) but still using the adjusted CPS-ASEC weights moves the counts closer to the administrative population.

Using IRS data and weights accounts for nonfiler income



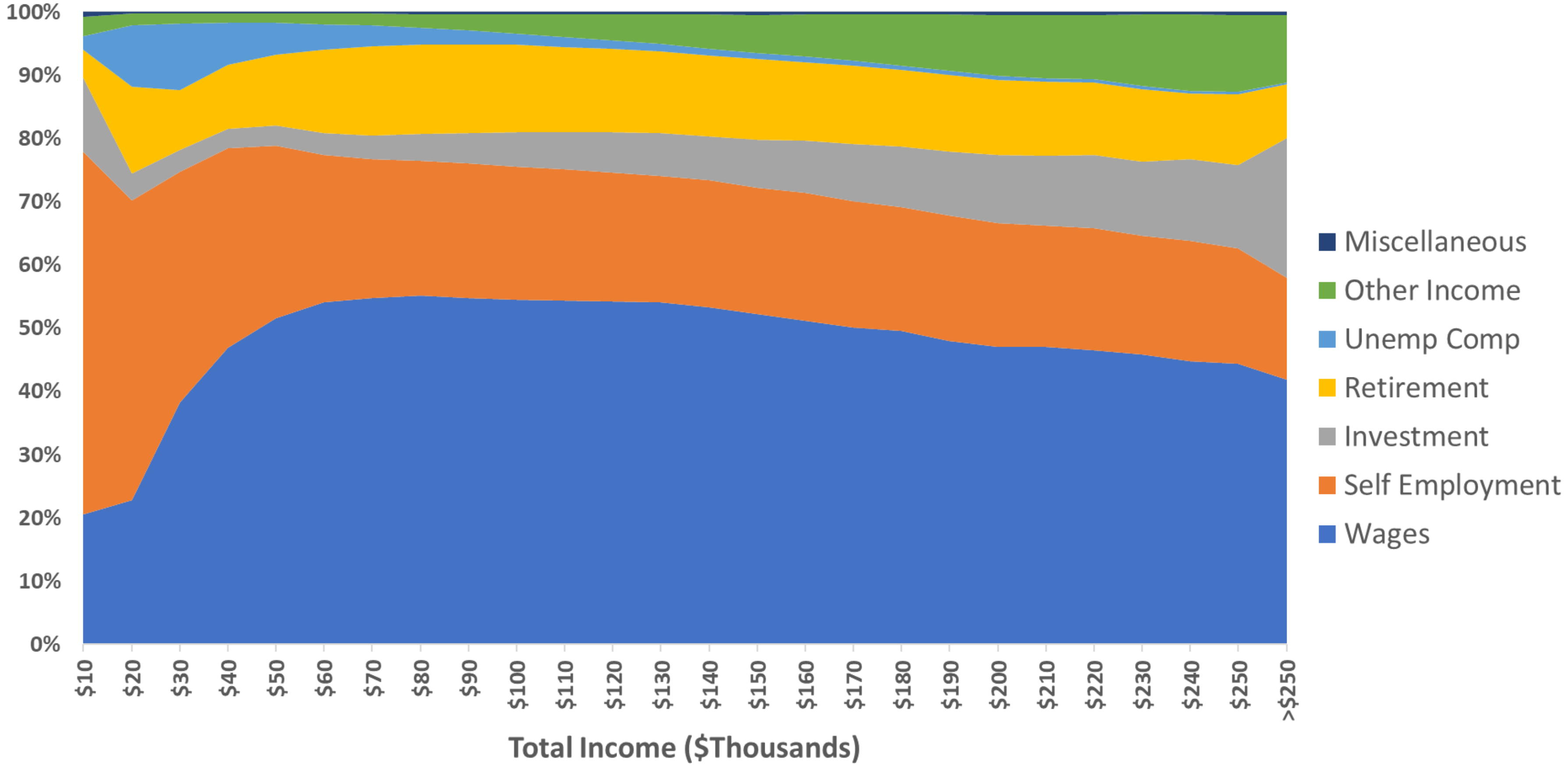
Now switching to the IRS weights while still using the comprehensive administrative data (with the SE imputation) brings the counts very close to the population counts (by design).

IRS weights raise total income to population level



- **CPS-ASEC linked to limited IRS data with imputations uses W-2s for wages and imputations for social security, pension, unemployment comp, and net SE earnings**
- **Comprehensive administrative data with SE imputation uses IRS third-party income except for net SE income**

S-E income is a much larger share of total income < $60K

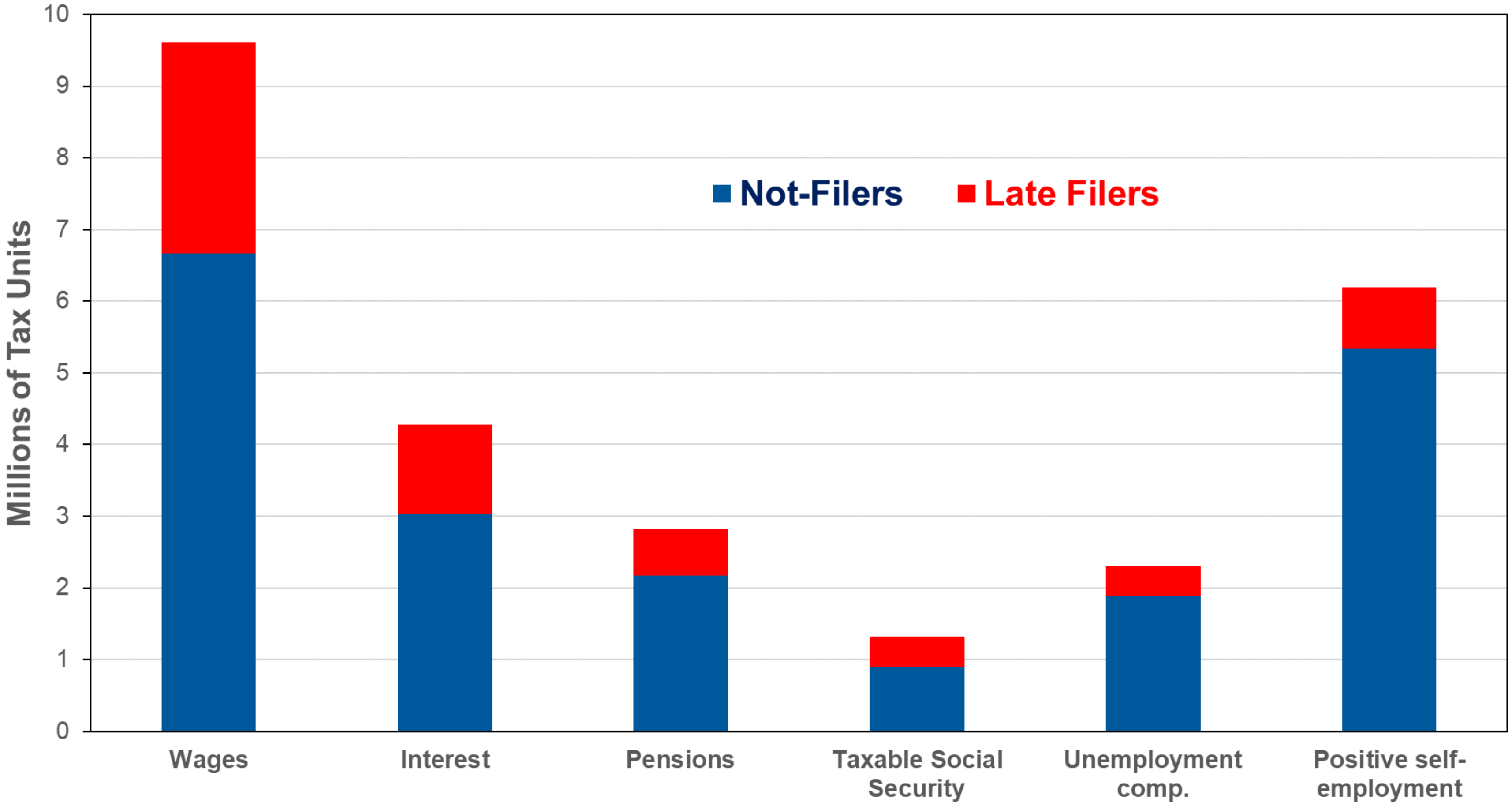


Late Filers

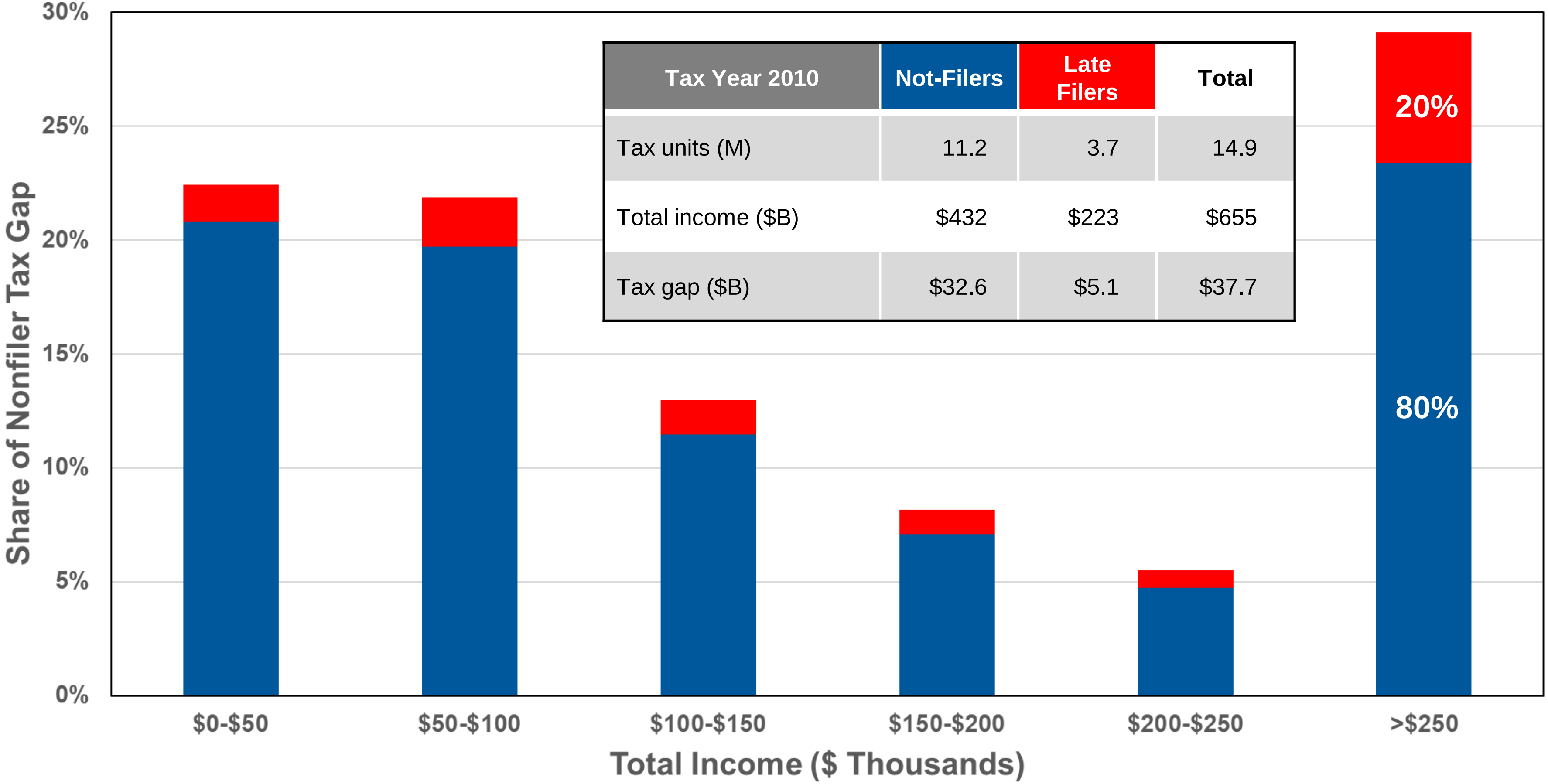
- For this analysis (given current one year data limitation), we define late filers as those who file after their filing deadline, but by December 31, 2011.

- Other returns are filed later than that, but those are represented among the not-filers.

Proportionally more wages & interest than other income among Late Filers



Late Filers account for 25% of Nonfilers, but only 13.5% of the Nonfiler Tax Gap



Tax Year 2010	Not-Filers	Late Filers	Total
Tax units (M)	11.2	3.7	14.9
Total income ($B)	$432	$223	$655
Tax gap ($B)	$32.6	$5.1	$37.7

Future Enhancements

- Explore options for addressing **downward biases**
  - Omissions from population?
  - Unaccounted for income reported to Census?
- Apply analysis to **multiple years**
- Compare results using **ACS** instead of CPS-ASEC
- Explore **drivers** of nonfiling via probit analysis

Using Uplift Modeling to Improve ACS Case Selection and Compliance Outcomes

Presented by: Jan Millard

Jan Millard, (IRS RAAS); Travis Whitfield, Sarah Smolenski, Michael Stavrianos, and Lauren Szczerbinski (ASR Analytics)

IRS TPC Joint Research Conference | June 24th, 2021

IRS

ACS Optimization and Notice Redesign Background

- ACS Collection notices like the LT11 and LT16 are issued to a subset of eligible taxpayers due to resource constraints. Optimizing ACS performance necessitates improvements in both:
  - **Notice Design** (i.e., increasing the effectiveness of the notice itself)
  - **Case Selection** (i.e., deciding which taxpayers should receive each notice)

- **Notice Design**
  - Randomized control trials (RCT) are run to test out different variants of notices, each including specific behavioral “nudges”
  - Redesign pilots have been run for the CP14, CP501/CP503, LT11, and LT16 notices

- **Case Selection**
  - Machine learning models are implemented to identify which taxpayers should receive a notice
  - The 2019 LT16 Model Test Pilot tested a case selection methodology using a non-uplift predictive model
    - Lessons learned from this pilot, especially the prevalence of refund offsets among the full-paying taxpayers, were incorporated into this follow-on uplift-model based effort

What is Uplift?

- Definition
  - Uplift is the incremental change in the likelihood an individual will take a specified action (or the change in magnitude of that action) in response to a treatment:

$$\boldsymbol{uplift} = \boldsymbol{P}(\boldsymbol{Y} = \mathbf{1}|\boldsymbol{T} = \mathbf{1}) - \boldsymbol{P}(\boldsymbol{Y} = \mathbf{1}|\boldsymbol{T} = \mathbf{0})$$

  - Where:

    T = treatment status (0 or 1)

    Y = outcome of interest (0 or 1)

- Constraint
  - Because a given individual cannot both receive a treatment and not receive a treatment, it's not possible to directly measure uplift
  - Instead, uplift is typically measured in aggregate with an RCT

Categories of Taxpayers

Uplift category matrix

Taxpayer behavior if sent notice

Taxpayer behavior if not sent notice	Payment	No Payment
Payment	**Sure Thing** Taxpayers who “self-cure” (i.e., full pay regardless of whether they receive a notice). Includes refund offsets.	**Do Not Disturb** Taxpayers for whom a notice has a negative impact on the likelihood to full pay. Also represents taxpayers who only call upon receipt of a notice.
No Payment	**Persuadable** Taxpayers who full pay only if they receive a notice.	**Lost Cause** Taxpayers who will not full pay regardless of notice receipt. Includes taxpayers who are not financially capable or unlikely to successfully receive the notice.

– The goal is to identify taxpayers in the *Persuadable* group, and only send notices to those taxpayers

– Much of the modeling effort is involved in identifying *Sure Thing, Do Not Disturb, or Lost Cause* taxpayers, in order to ensure they are <u>not</u> sent a notice

Uplift Modeling Approaches

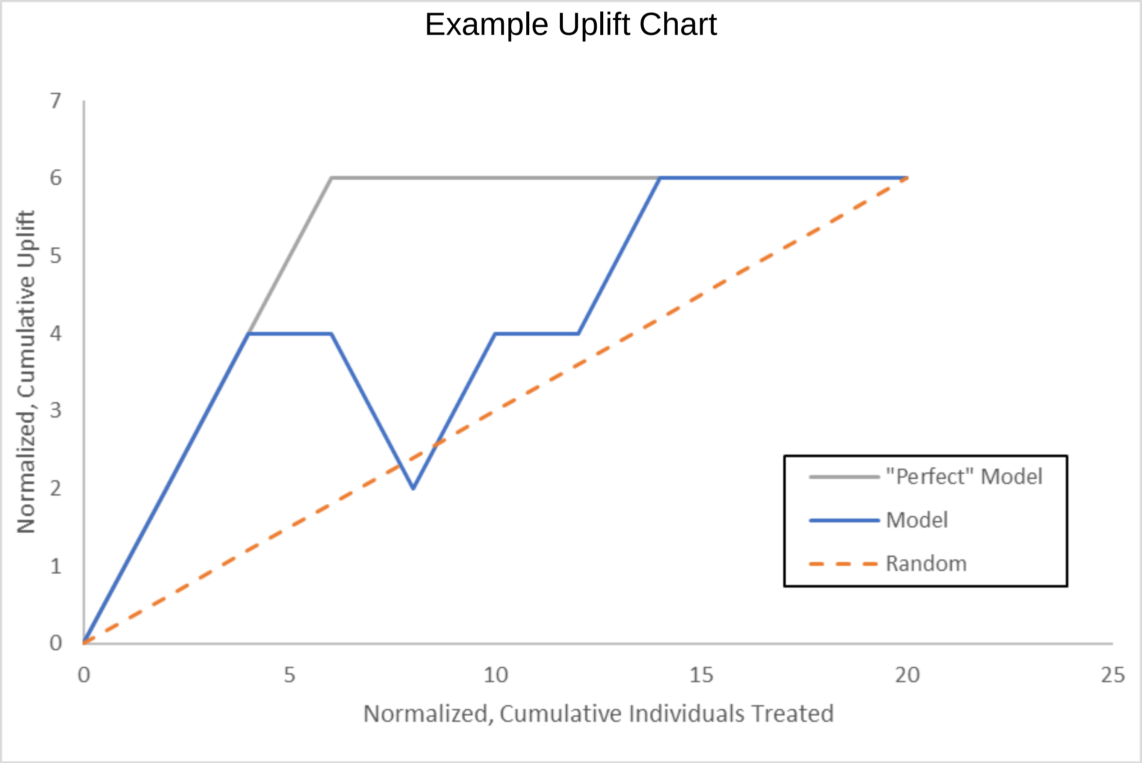
- **Two-Model**
  - Two separate models are trained on the treatment and non-treatment groups
  - Predicted uplift is the difference between the predictions of the two models

- **Single-Model**
  - One model is trained on combined dataset of treatment and non-treatment groups, where treatment status is a feature in the model
  - Predicted uplift is the difference between prediction of the model when T = 1 and when T = 0

- **Class-Transformation**
  - New target variable, Z, is created such that Z = 1 when T = 1, Y = 1 and when T = 0, Y = 0, and Z = 0 otherwise
  - Single model is trained to predict the new target variable, then predicted uplift is calculated:

$$\boldsymbol{uplift = 2 * P(Z = 1) - 1}$$

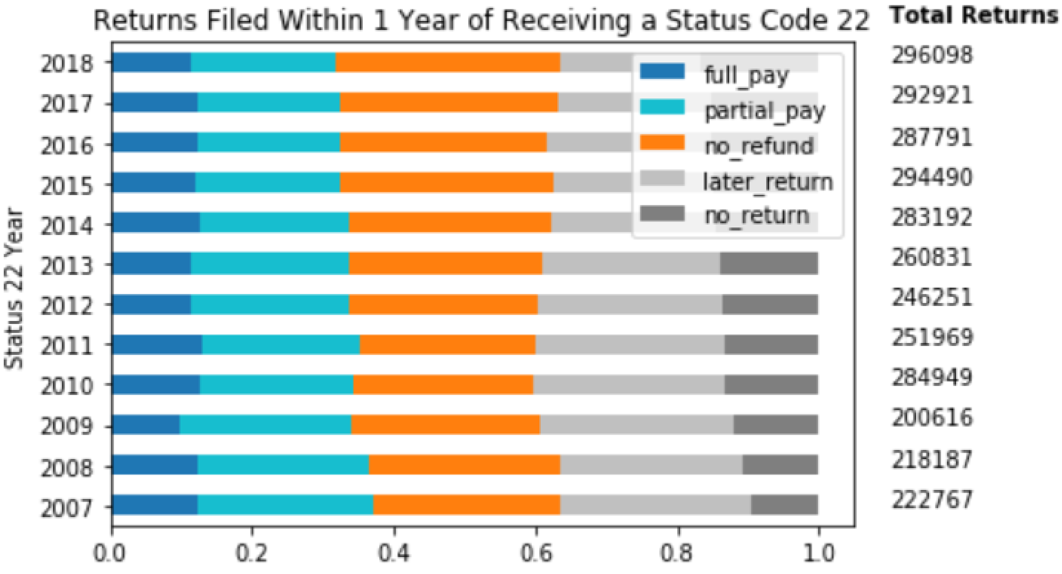
Uplift Model Evaluation



- Uplift models are evaluated by plotting the normalized cumulative uplift vs. the normalized cumulative individuals treated
- The area under this curve (Uplift AUC) can be used as a single metric to compare different models (corresponding to the same dataset)
  - The Uplift AUC can be normalized by dividing it by the area under the “Perfect” Model curve

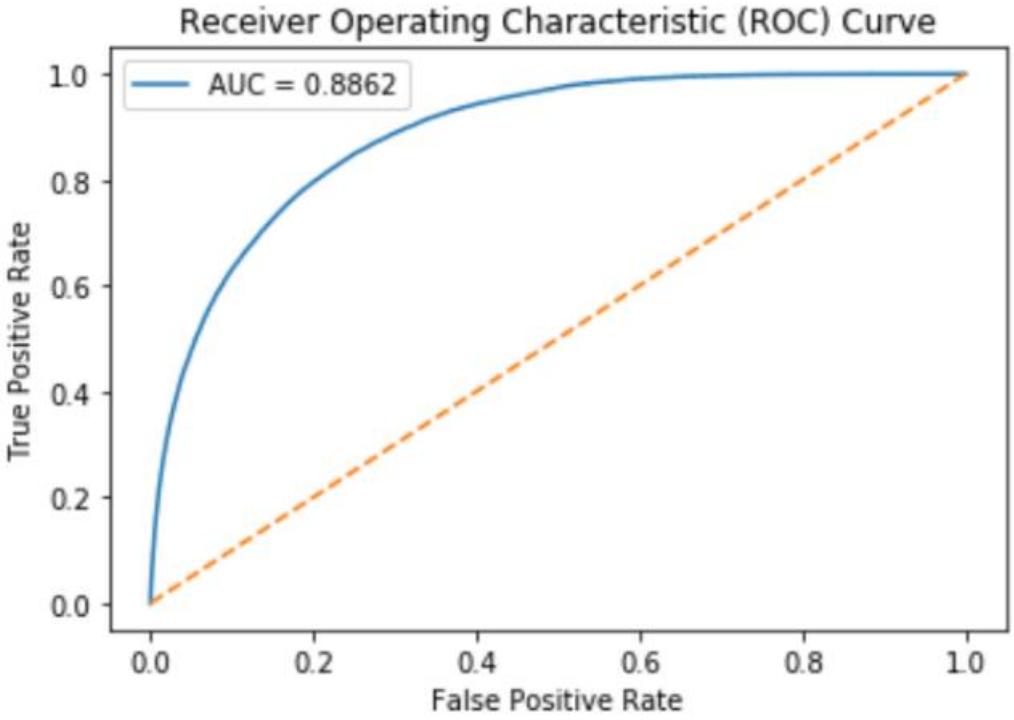
Refund Offset Exploratory Analysis

- Approximately 35% of taxpayers in ACS file a tax return with a refund within one year of entry, and about 10% file a refund large enough to result in the full payment of their balance due
- Successfully predicting which taxpayers will fall into that 10% would greatly improve case selection, allowing the IRS to focus limited resources on other taxpayers who are more likely to benefit from receiving an LT11 notice



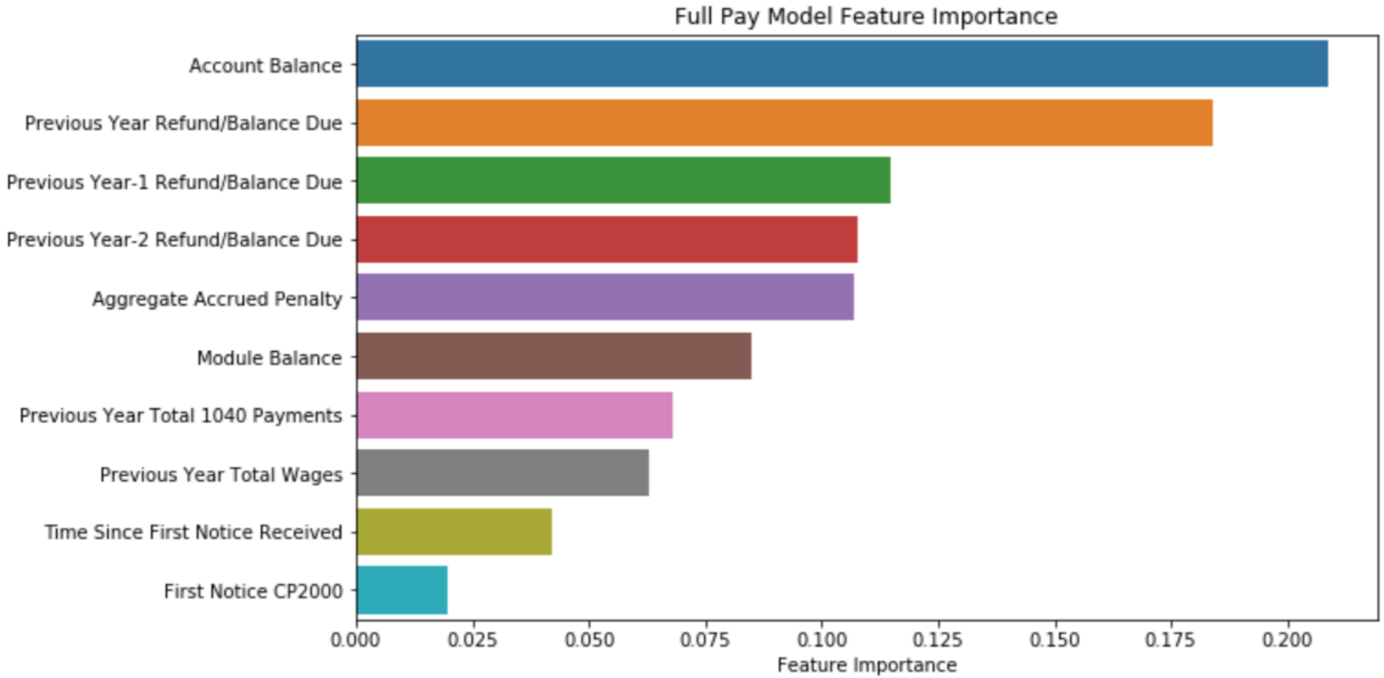
Refund Offset Model Results

- The resulting full payment refund offset prediction model was highly effective

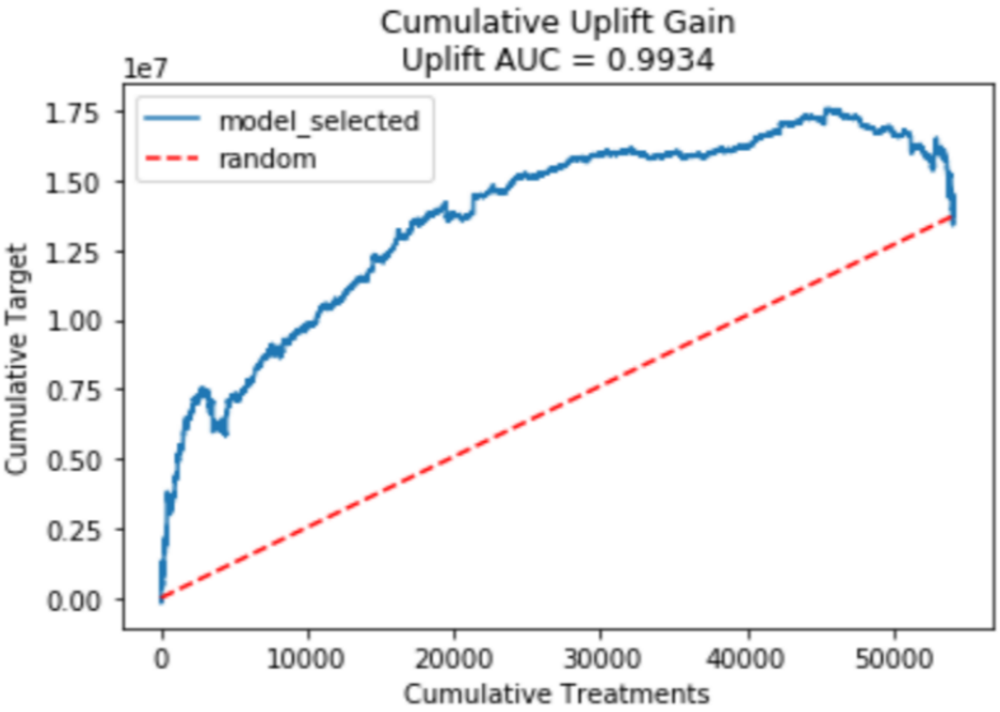


Refund Offset Model Results

- As expected, total account balance and historical tax return refund / balance due amounts were the features with the highest predictive power



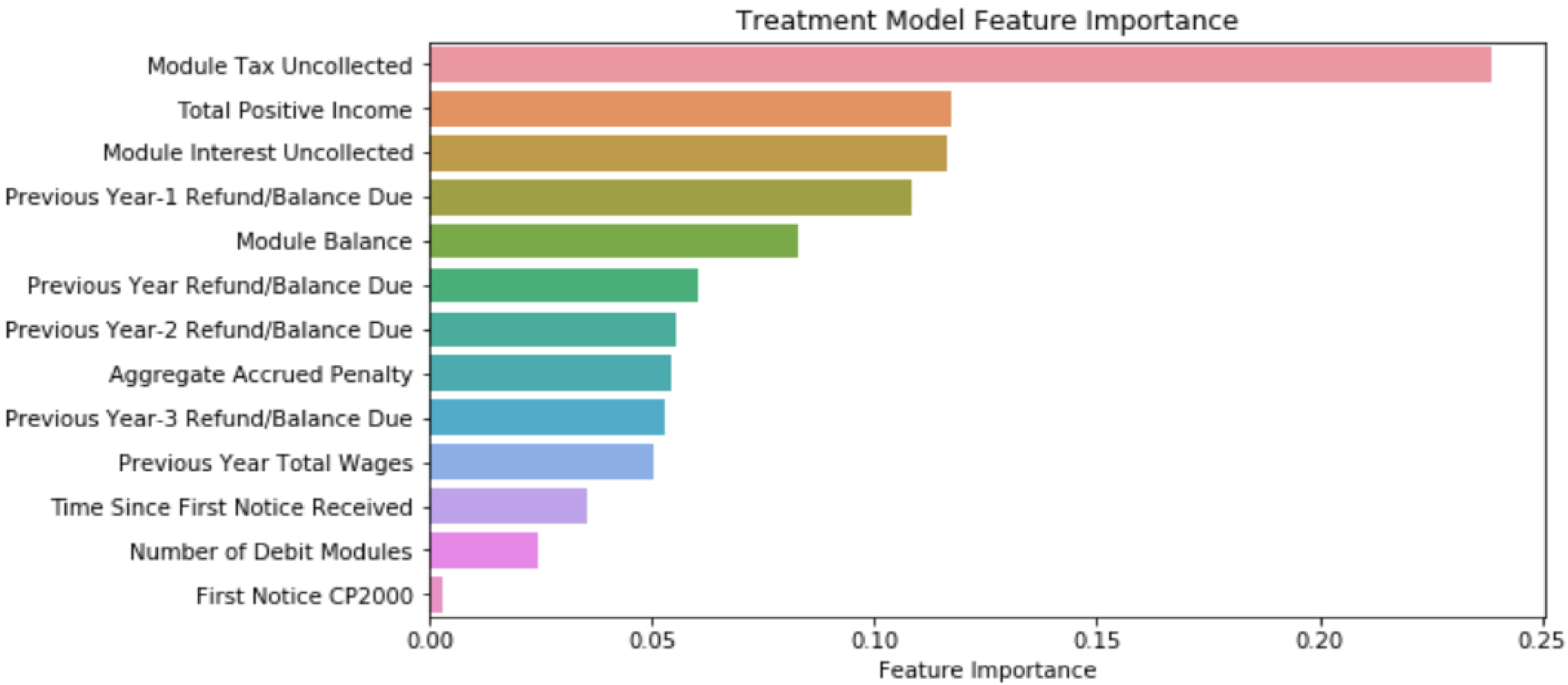
LT11 Total Dollars Uplift Model Results



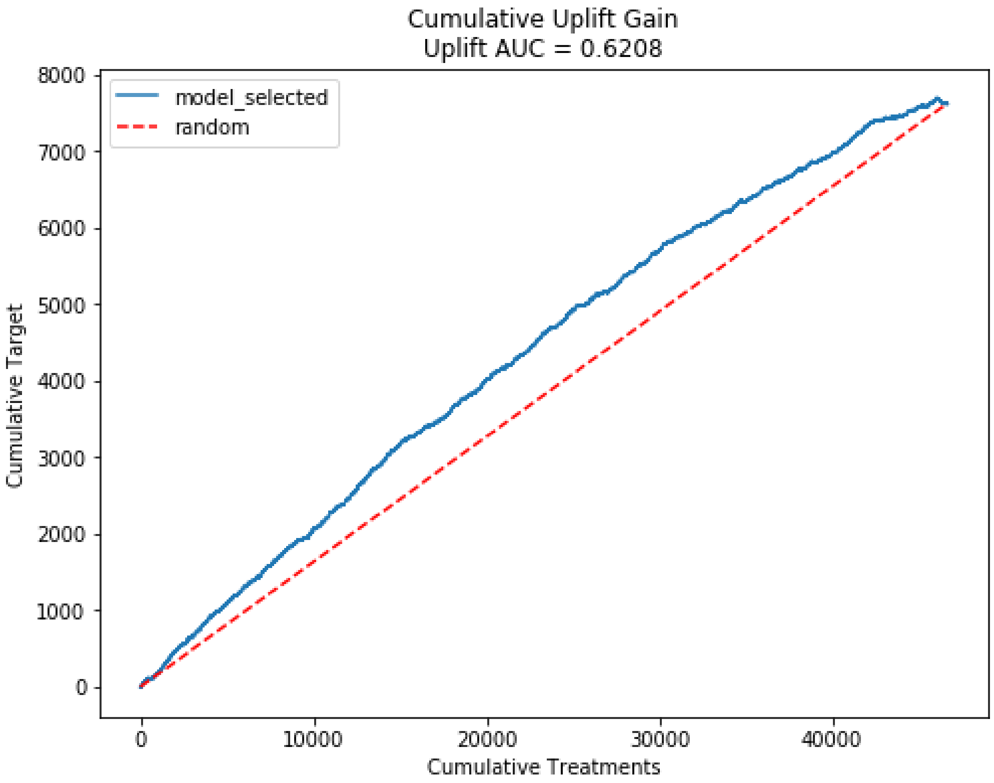
- This uplift model predicted the uplift in total dollars collected within a 365-day period following receipt of an LT11 notice
  - The results show the large improvement in dollars collected that can be achieved by sending notices to a fraction of the total eligible recipients

LT11 Total Dollars Uplift Model Results

- Many of the features of the total dollars collected model are the same as those for the refund offset model, indicating that refund offsets play an important part



LT11 Phone Calls Uplift Model Results



- This uplift model predicted the uplift in taxpayers who call the IRS within a 365-day period following receipt of an LT11 notice
  - There appears to be less predictive power in the LT11 Phone Call model than in the LT11 Total Dollars Collected model, but the results do indicate that the model can help identify which taxpayers are more likely to call the IRS after receiving their LT11 notice

Conclusions

- Uplift modeling is a promising approach to optimizing ACS case selection processes
  - A pilot test to evaluate this methodology will be run in the coming months
- The refund offset model is a promising approach to identifying taxpayers who are likely to self-cure
  - This modeling can be done without the need for a dedicated RCT, which expands the possible applications where this sort of approach can be taken

Enhancing Return Risk Assessment for Examination: Recent DIF (Discriminant Function) Model Updates

Getaneh Yismaw, Drew Johns, Taukir Hussain, Jonathan Creem, Mary-Helen Risler
*(IRS, Research, Applied Analytics & Statistics) **

*The authors are analysts and their manager in the Knowledge Development & Application Division, Compliance Modeling Lab. The views expressed in this paper do not necessarily represent the views of the Department of the Treasury or the Internal Revenue Service.

IRS/TPC Joint Research Conference on Tax Administration

June 24, 2021

Outline

- Introduction
- DIF Background and Overview
- DIF Methodology
- Difference between DIF and Standard Linear Discriminant Analysis
- Interpretation and Application
- Model Testing (Evaluation) and Selection
- Continuous Model Improvement
- Examination Results from Updated DIF Models
- Exploring Enhancements to Return-Level Risk Modeling
- Future Work

Introduction

- To fulfil its mission, the IRS works to promote voluntary compliance:
  - ❑ Providing taxpayer services and education
  - ❑ Strategically enforcing the law
- IRS employs a balanced compliance enforcement strategy:
  - ❑ Enforcement presence across all types of tax returns
  - ❑ Returns that are most likely to have significantly underreported taxes are audited at a relatively higher rate
- A variety of compliance enforcement strategies and tools commensurate with the nature of the noncompliance:
  - ❑ Math error corrections during return processing
  - ❑ Matching return information to third-party information
  - ❑ Examinations
    - ✓ Campus correspondence examinations
    - ✓ Office examinations
    - ✓ Field examinations

Introduction, cont'd

- The IRS has many strategies and methods for selecting returns for an examination:
  - ❑ Information from whistle blower
  - ❑ Referrals
  - ❑ Compliance risk models

- IRS uses DIF models as one method to assess compliance risk for examinations:
  - ❑ Individual income tax (the focus of this presentation)
  - ❑ Small corporation income tax,
  - ❑ S-corporation and
  - ❑ Partnership returns

DIF Background and Overview

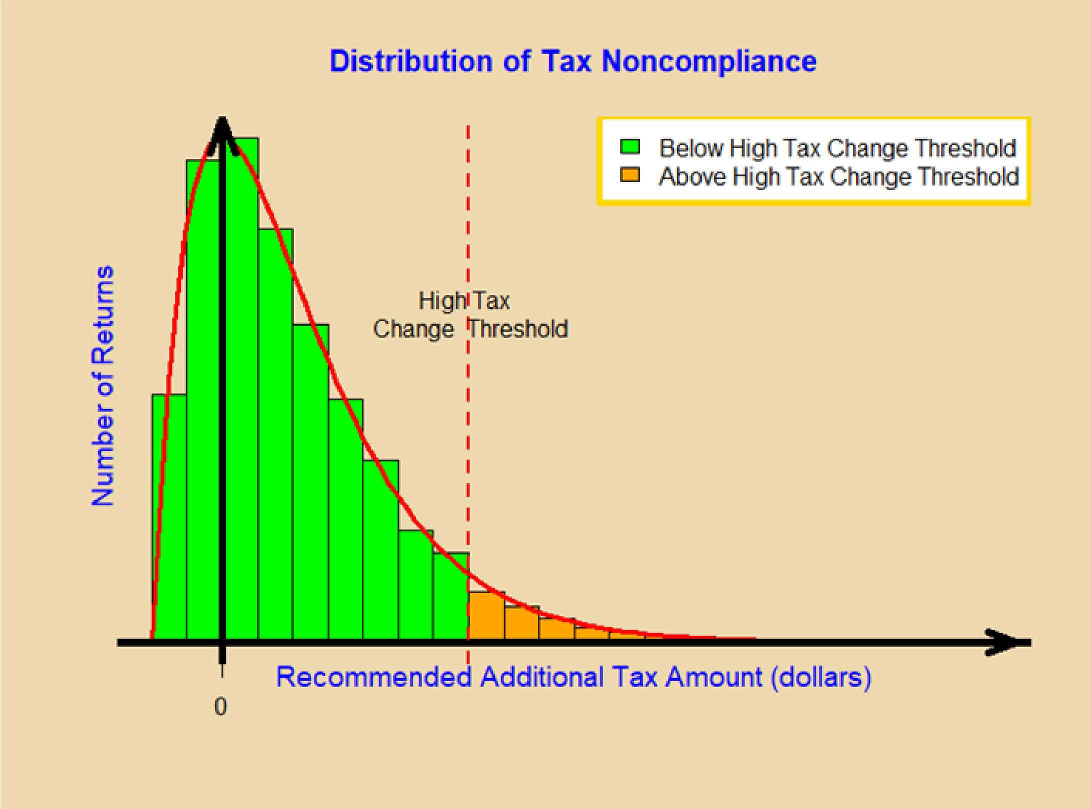
- DIF is a supervised machine learning technique that predicts the likelihood of a significant tax change at the tax return level
- Developed by IRS in the 1960s based on Fisher's linear discriminant analysis
- Refined and improved over the years
- The effectiveness and robustness of DIF has been widely recognized:
  - ❑ Internal agency research and reviews of actual field examination results
  - ❑ Outside independent technical reviews
- DIF models are developed using tax return information and examination results from statistically representative samples of tax returns
  - ❑ National Research Program (NRP), and previously the Taxpayer Compliance Measurement Program (TCMP)
- DIF models are implemented as part of tax return processing
  - ❑ Score becomes part of administrative data systems for downstream applications
- DIF remains one of the main sources of return selection for field and office-based audits

DIF Methodology

- Linear discriminant analysis (LDA) is a form of a multivariate statistical technique that is designed to classify observations into two or more groups based on a characteristic of interest
- DIF is used to develop predictive models to separate returns into two groups: returns that have a high likelihood of a high tax change from those that do not
- Goal is to select returns from the extreme tail of the misreporting distribution
- Based on data driven analysis, a "high tax change threshold" level is set for each activity code
- The IRS stratifies returns into mutually exclusive groups of classes, known as, activity codes.

❑ Complexity, nature, and level of underreporting vary by activity code

- "High tax change threshold" levels may vary across activity codes
- The basic approach is to identify predictor variables, that in combination, do the best job of predicting returns with the largest amount of underreported tax



Difference between DIF and Standard LDA

- In DIF, model parameters are estimated, just as in any other LDA technique by inverting the pooled variance-covariance matrix for the two groups

- DIF is a modified application of LDA in which predictor variables are transformed into a set of discrete likelihood ratios

- Binning (intervalization) and likelihood ratio transformation:
  - ❑ Predictor variables are pre-processed into wide ordinal intervals
  - ❑ For each predictor variable a likelihood ratio is computed for each interval
  - ❑ Likelihood ratios replace all potential predictor variables

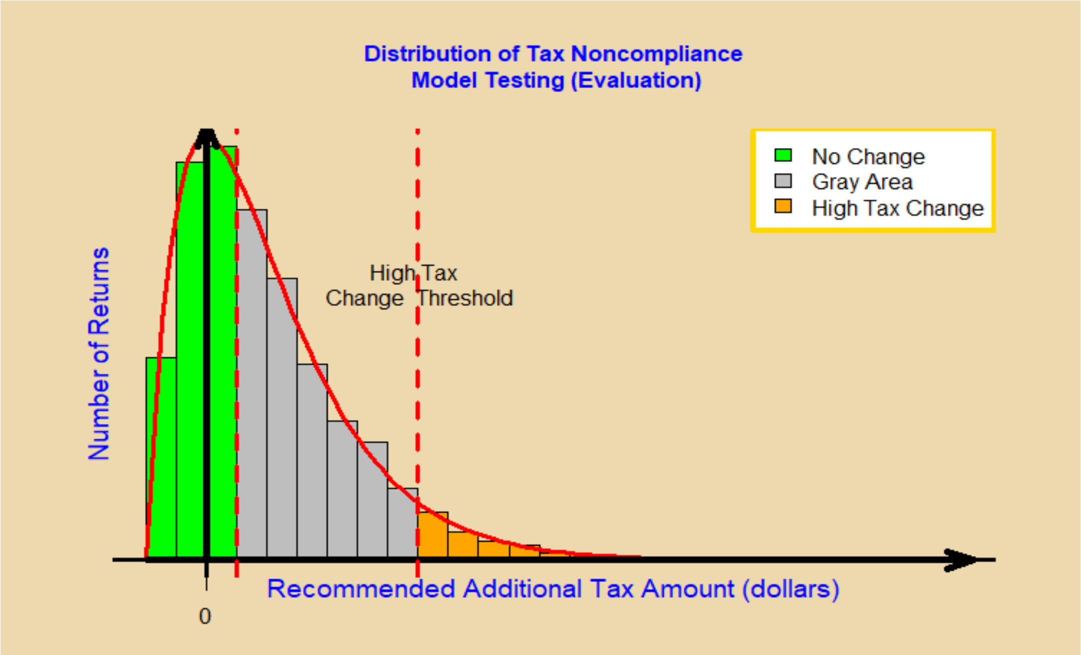
- Value of binning and transformation of inputs into likelihood ratios:
  - ❑ Well suited for data that is as complex as tax return data
  - ❑ Handles line items that have large numbers of zero entries and extreme outliers
  - ❑ Effective at reducing noisy and uninformative predictor variables
  - ❑ Contributes to the robustness of DIF

Interpretation and Application

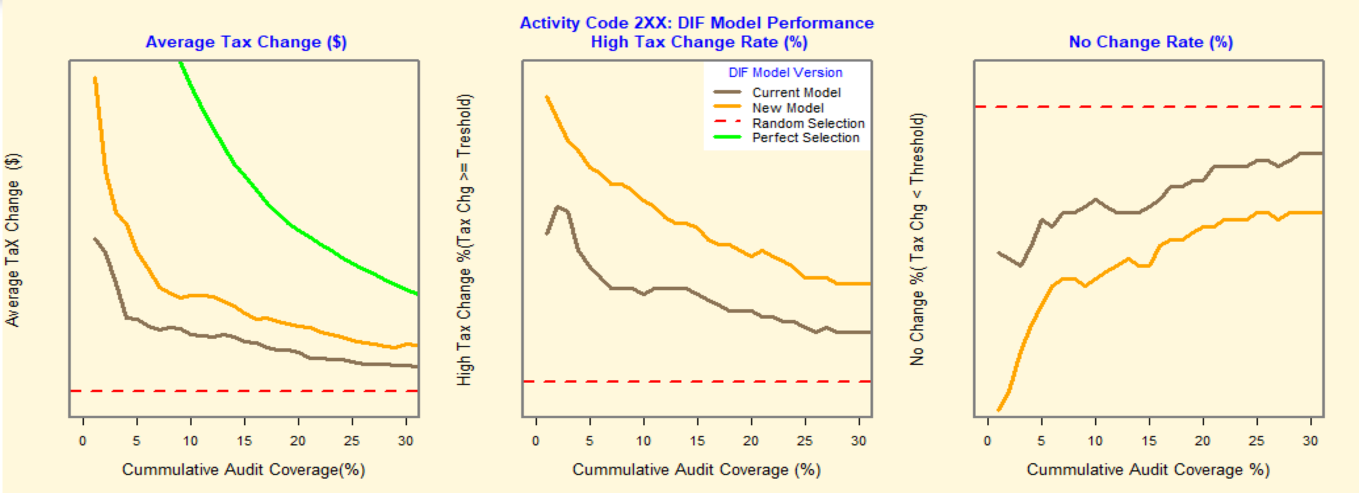
- Final DIF scores are scaled (normalized) to fall within a specified range
- Returns assigned a high DIF score by the models have a higher likelihood of significant *overall* tax change
- Relative difference in the magnitude of two DIF scores is not necessarily reflective of the relative difference in their compliance risk
- Case selection for examination is designed to be “top-down” from high to low ranking by DIF score
- It is the rank of a return’s DIF score within an activity code that matters for risk ranking and application
- Separate DIF models are developed for each activity code
- The distribution of DIF scores is not necessarily the same across activity codes
- DIF scores cannot be compared across activity codes or between different model versions (different tax filing years) within an activity code

Model Testing (Evaluation) and Selection

- Model Testing (Evaluation) and Selection Criteria
  - ❑ Average Tax Change
  - ❑ "No Change" Rate
  - ❑ High Tax Change (or "Hit") Rate
- Average tax change ($): the overall average examination recommended additional tax
- "No Change" Rate (%): proportion of returns with tax change less than the "no-change" threshold (green region)
- High Tax Change rate (%): the share of returns expected to result in "high tax change"



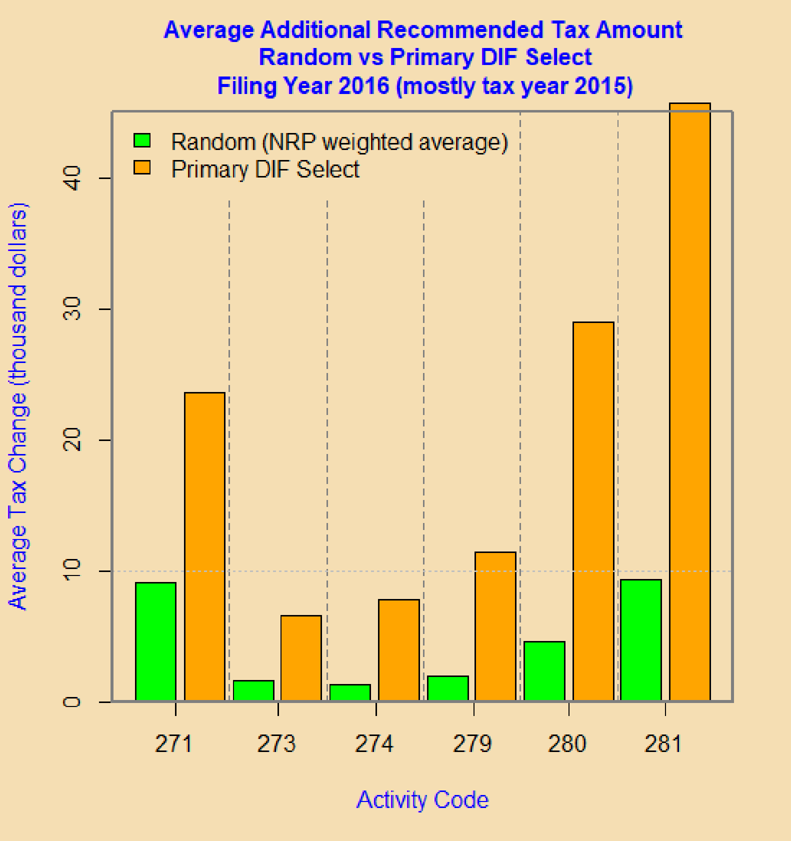
Model Testing (Evaluation) and Selection: Example of DIF Model Evaluation



- Models are evaluated at cumulative coverage levels, say at 1%, 2%, …, 20%
- DIF models are designed to be effective at the upper tail of the misreporting distribution
- New competitor model significantly outperforms the existing model across all three model evaluation criteria using a test data set
- There are circumstances where a model shows significant improvement in one criterion but not in another

Model Testing (Evaluation) and Selection: DIF Model Effectiveness

- Average additional recommended tax change: primary DIF select vs population average estimates
  - ❑ Primary DIF select returns filed in 2016 and closed by March 2021
  - ❑ If there were no risk models, the expected average tax change from random selection is the weighted NRP estimate



Continuous Model Improvement

- DIF model updates that have been made using NRP TY2006+ data
  - Beginning with tax year 2006, NRP restarted its individual income tax studies with a smaller annual sample design which required combining multiple study years
  - Because of the time it takes to complete tax audits, the first DIF model development took place in 2011-2012 for implementation in filing year 2013

Table : DIF Model Development and Implementation, Filing Years 2013-2022

Implementation Filing Year	NRP Data Used	Number of Activity Codes	
		Developed	Implemented
2013	TY2006-2008	5	2
2014	TY2006-2009	5	4
2016	TY2006-2010	7	6
2019	TY2006-2013	6	5
2022**	TY2009-2015	9	6

** For filing year 2022, the IRS is increasing the number of individual income tax activity codes from 12 to 14. Activity code 281 will be divided into three new activity codes (each with its own new DIF scoring model). The new activity codes are as follows:

282: Total positive income >= $1,000,000 and < $5,000,000

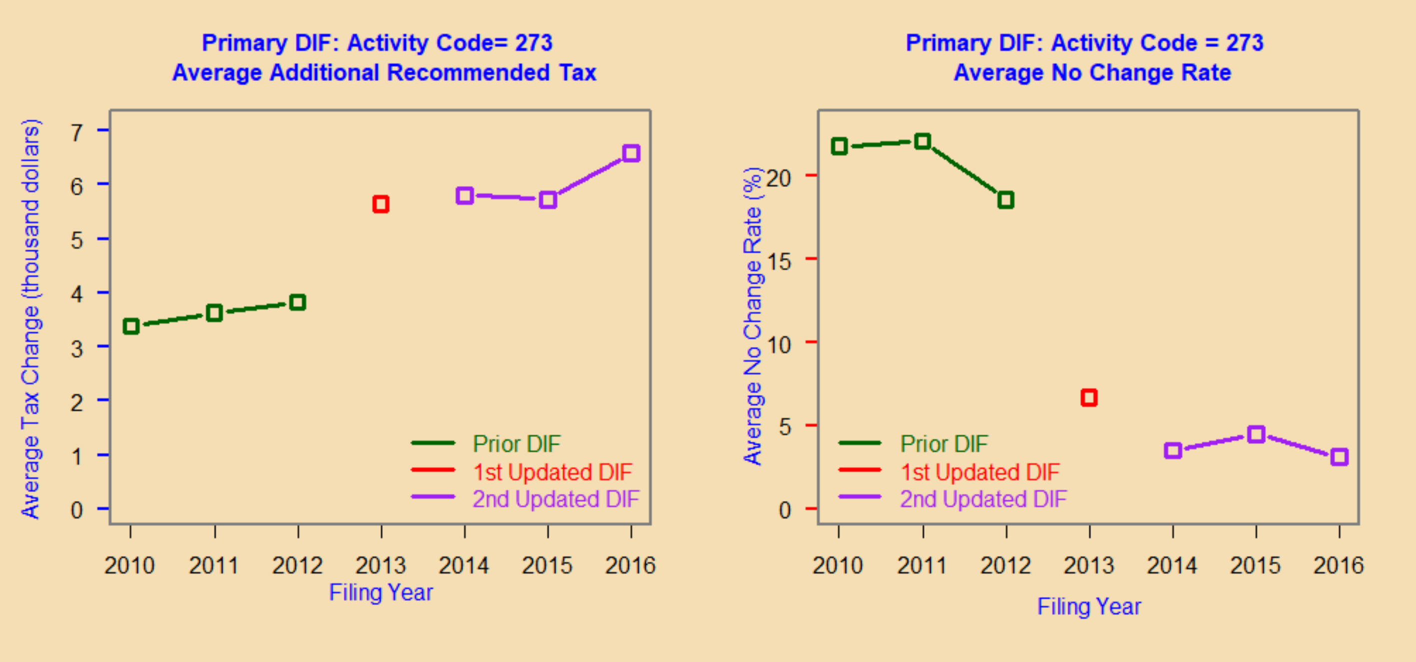
283: Total positive income >= $5,000,000 and < $10,000,000

284: Total positive income >=$10,000,000

Examination Results from Updated DIF Models

- The following five graphs demonstrate the effects of model improvements made in the first three rounds of DIF updates for five activity codes
- These five activity codes were updated twice over this time period
- Trends in average tax change and no-change rates from filing-years 2010-2016 for returns closed by March 2021

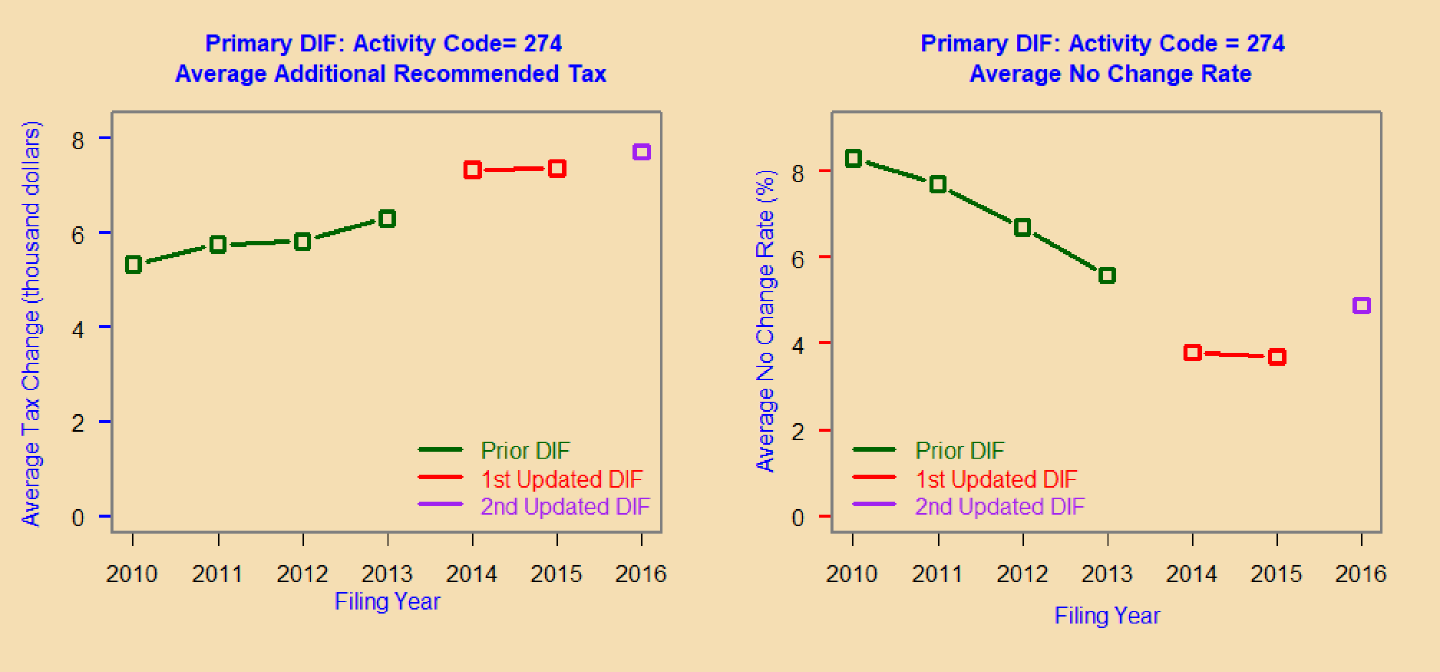
Examination Results from Updated DIF Models: Activity Code 273



- Activity code 273 includes non-business returns with total positive income under $200,000; with no Schedule C/F; with Schedule E or Form 2106 present

*Source: Audit Information Management System (AIMS) , cases closed by March 2021

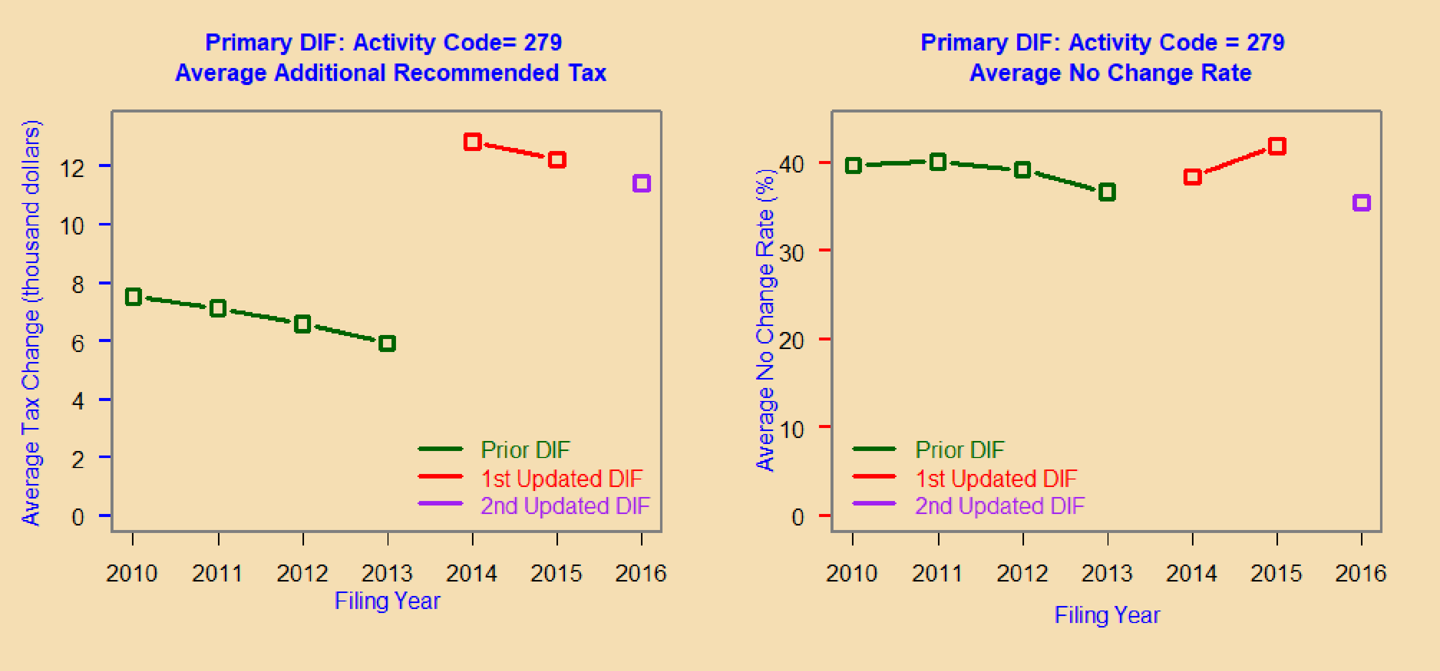
Examination Results from Updated DIF Models: Activity Code 274



- Activity code 274 (returns with total positive income < $200,000; with Schedule C; total gross receipts < $25,000)

*Source: Audit Information Management System (AIMS) , cases closed by March 2021

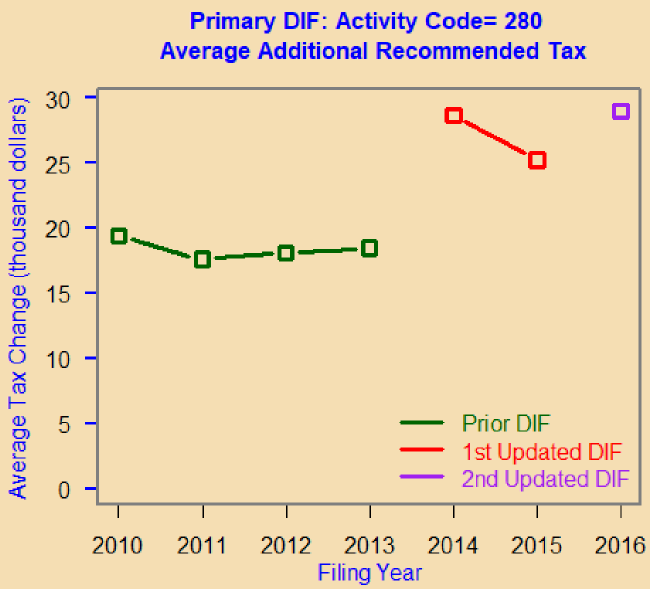
Examination Results from Updated DIF Models: Activity Code 279

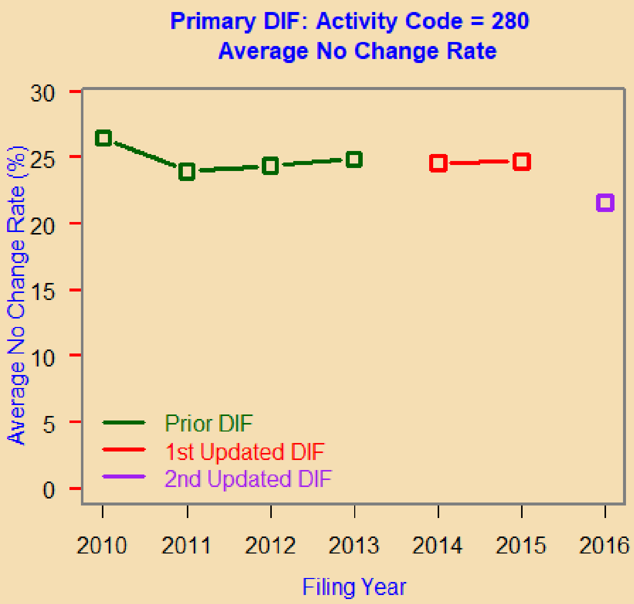


- Activity code 279 includes non-business returns with total positive income at or above $200,000 but less than $1 million; with no Schedule C/F; with Schedule E or Form 2106 present

*Source: Audit Information Management System (AIMS) , cases closed by March 2021

Examination Results from Updated DIF Models: Activity Code 280

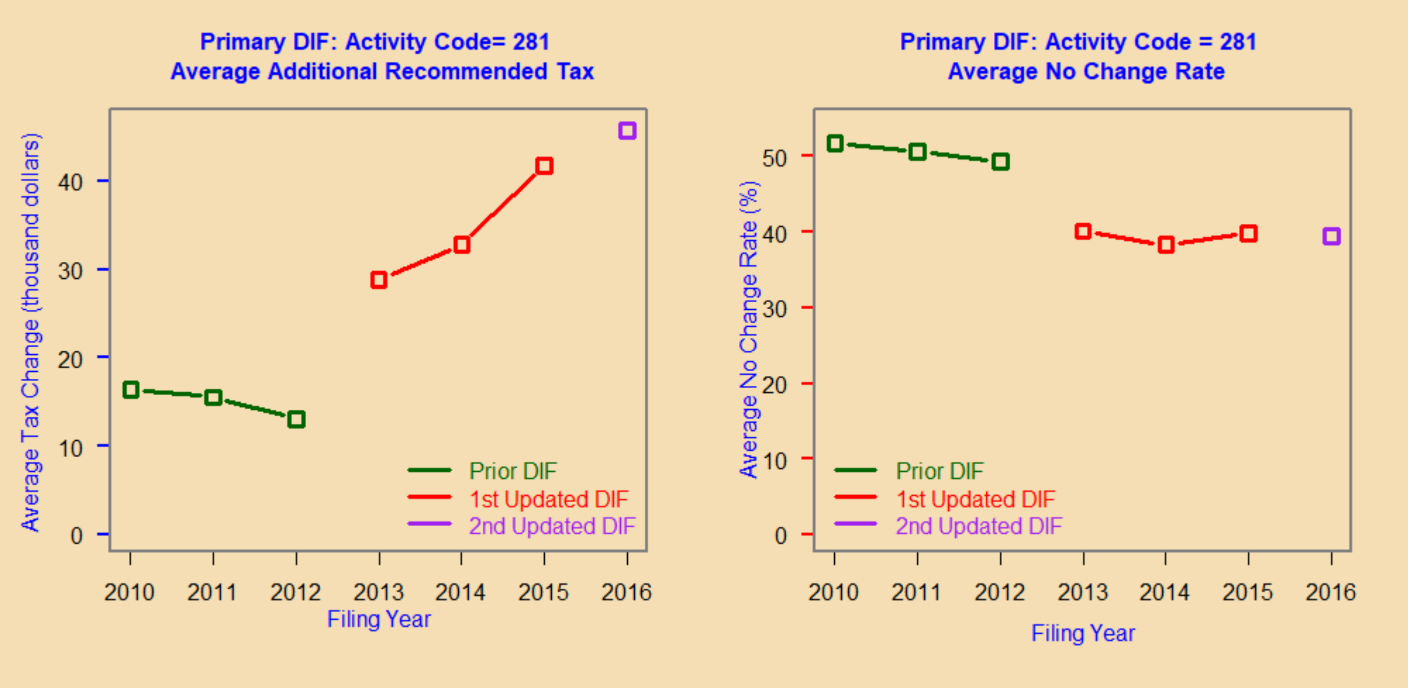




- Activity code 280 returns are business returns with $200,000 <= total positive income < $ 1 million; with Schedule C/F

*Source: Audit Information Management System (AIMS) , cases closed by March 2021

Examination Results from Updated DIF Models: Activity Code 281



- Activity code 281: returns with total positive income >= $ 1 million

*Source: Audit Information Management System (AIMS) , cases closed by March 2021

Exploring Enhancements to Return-Level Risk Modeling

- Research into improving return level risk modelling generally falls into two categories: exploring alternative analytical techniques and exploring additional potential predictor variable
- Prior research into alternative techniques has included the following:
  - ❑ Various two-stage (probability * regression) modeling efforts
  - ❑ Early versions of neural networks and Classification and Regression Trees (CART)
  - ❑ Enhancements to return-level risk modeling: Issue level models
- Recently completed contract to develop machine learning algorithms for two activity codes:
  - ❑ Various parametric regression models, tree-based models, and other classification algorithms
  - ❑ Various model pipeline steps
  - ❑ In general, slightly better no-change rate to current DIF but lower tax-change
- Internal research
  - ❑ Random forest
  - ❑ Comparable performance to DIF when evaluated with out-of-sample data
- Consistent with early research, recent results show DIF is effective and robust

Future Work

- Continue updating DIF models as new data become available
- Continue exploring alternative techniques
- Expand the universe of potential predictor variables
- The research to date into alternative techniques suggests that no one analytical technique yields models that are vastly superior to all others on all relevant criteria
- Explore ways that models developed using different analytical techniques might complement one another as part of a process for return selection that could improve the overall strategy for examination return selection

Developments in Technology and Analytics

Brian Erard

Common theme across papers

- Making the best possible use of available data to assess and address compliance issues

Study by Tom, Pat, Mark, Alan & Maggie

- Remarkable achievement to link detailed IRS administrative data with CPS-ASEC data
- Results so far help to clarify potential nature and size of individual nonfiler population
- Ongoing work promises to bring further insights
- Findings about CPS-ASEC data quality raise potential concerns about the quality of poverty estimates as well as the data source's suitability for many existing and future research initiatives

Which nonfilers might be missing from IRS admin data?

- Nonfilers who do not have interest-bearing bank accounts or other income sources reported on 3rd party information returns
  - Some undocumented immigrants and others who are paid under the table
  - Some self-employed taxpayers
- Individuals who provide phony TINs to third-parties

Which nonfilers might be missing from CPS-ASEC?

<table>
<tr><th>Hard to survey populations</th><th>Reasons</th></tr>
<tr><td>

Young children
Highly mobile persons
Racial and ethnic minorities
Non-English speakers
Low income persons
Persons experiencing homelessness
Undocumented immigrants
Persons who distrust the government
LGBTQ persons
Persons with mental or physical disabilities
Persons who do not live in traditional housing

</td><td>
Hard to locate

Residence not in sampling frame
Hard to interview
language barriers
Literacy
lack of Internet

Hard to contact

Highly mobile
Transient/homeless,
Gated communities/Celebrities

Hard to persuade

Suspicious of government
Low civic engagement
Fears about deportation

</td></tr>
</table>

Adapted from Chapin et al. (2018)

Re-weighting leaves out some nonfilers

Admin Data	CPS-ASEC Data		
	Not Present	*Present/ No PIK*	*Present/ PIK*
Not Present	A	B	C
Present but excluded	D	E	F
Present & included	G	H	I

Group A – Not counted / hard to address
Groups B & C – Not counted / potentially addressable
Groups D, E, & F – Not counted / hard to address
Groups G & H – Counted indirectly via re-weighting
Group I – Counted directly

Attempt to gauge how many nonfilers might be missing in admin data

- How many filers with no 3rd party information returns?
- How many CPS-ASEC individuals who exclusively report income from sources unlikely to be reported by a 3rd party?

Re-weighting

- How well does re-weighted Group I represent Groups G, H, and I combined?
  - Current re-weighting approach does seem sensible.
  - However, Bond et al. (2014) found that minorities, residents of group quarters, immigrants, recent movers, low-income individuals, and non-employed individuals are less likely to receive a PIK
    - These factors not included in current probit-based weighting approach
    - Many of these factors may not be available for use in admin. data.

Quality of admin. income measures

- Imputed self-employment income for some sample members
  - Could improve further by including DCE estimates
  - Might replace mean prediction with randomized draw from distribution
- Might impute Sch. E, cap gains, tip income, other sources not routinely reported 3rd party returns

High-income nonfilers

- Should investigate to make sure they really account for such a large share of nonfiler tax gap
  - PIK mismatches?
  - Very late filing?
  - Foreign tax credits & other credits?
  - Evidence from recent high income nonfiler enforcement activities?

Uplift study by Jan, Mike, Lauren & Travis

- Nice Features
  - Careful modeling process
    - Consideration of target variable and time window for response
  - Rigorous feature selection, training, tuning, and testing process for alternative models
  - Scoring process explicitly accounts for resource constraint
  - Scoring process evaluates success on basis of uplift, thereby discounting cases where a taxpayer was likely to “self-cure” even in the absence of the treatment

Taxpayers are heterogeneous

- They differ in terms of:
  - Likely response to a given compliance treatment
  - Revenue impact of their response
  - The burden they experience from the treatment
  - The IRS resource costs associated with their treatment
- So, one wants to apply the right compliance treatment to the right taxpayer
  - (including perhaps not treating some taxpayers at all)

The IRS is budget-constrained

- It does not have the resources to treat all taxpayers
  - Consequently, it is desirable from a cost-efficiency perspective to:
    - Prioritize taxpayers for a given treatment in accordance with the highest expected "bang for the buck"
    - Allocate resources across compliance programs according to where they are most productive on the margin.
- Even if IRS was not budget constrained, it would be wasteful to treat everyone

Potential impact of a notice

Not Treated	Treated	
	Pay	No Pay
Pay	"Sure Thing"	"Do Not Disturb"
No Pay	"Persuadable"	"Lost Cause"

So who should be treated?

- From a cost-effectiveness perspective, one should …
  - Avoid "Do Not Disturb" taxpayers
  - Avoid "Sure Thing" taxpayers
  - Seek out "Persuadable" taxpayers
    - Especially if they would contribute a lot of revenue relative to resource cost of treatment
    - (But avoid persuadable taxpayers experiencing financial hardship)
- What about "Lost Cause" taxpayers? …

"Lost Cause Taxpayers"

- Some compliance treatments are required in order for IRS to obtain legal authority to take certain compliance actions"
  - An LT11 notice must be issued before wages can be garnished or property can be seized
  - A substitute for return must be sent to a nonfiler taxes can be assessed based on third-party information
  - A CP2000 or Letter 516 must be issued before additional tax can be assessed based on apparent understatements of certain income sources or questionable reported offsets.
- It often DOES make sense to apply such treatments to taxpayers who are unlikely to respond.

Concerns beyond cost-effectiveness

- Some would question whether it is fair to avoid treating certain noncompliant taxpayers just because they are somewhat more costly to treat.
- In some cases, a strict cost-effectiveness strategy may have unintended consequences
  - e.g., incentivize taxpayers to become more costly to treat

DIF study – Getaneh, Drew, Taukir, Jonathan & Mary-Helen

- When first introduced, DIF program was way ahead of its time:
  - Large scale random audit programs still rather novel
  - Segmentation of population into different exam classes
    - Attempt to address systematic differences in compliance patterns across classes
  - Careful consideration of groupings of different compliance outcomes
  - Construction of explanatory variables based on binning and "likelihood ratio"
  - Systematic variable selection algorithm
  - Model refinement by discarding lowest-scored returns and re-estimating on remaining sample

Evolution of DIF program

- Many features of modeling process have stayed more or less the same over time
- One significant improvement is that alternative models are now evaluated based on out-of-sample performance.
- Another improvement is the skill sets of the researchers who work on model refinements, evaluation, and testing (part of a trend throughout RAAS)
- However, some of the model evaluation approaches discussed in the paper are not ideal
  - Comparison against random audits sets a very low bar
  - The trendline analysis, while perhaps suggestive, is not definitive
    - Might consider field testing new vs. old models (or comparing predictions using models based on one year of data on subsequent data years)

Updating models

- Would be interesting to explore updating models on a more frequent basis using a combination of new NRP results and results from some earlier NRP studies

Legacy activity codes

- Existing activity codes may not be the best basis for segmenting taxpayers into groups for analysis
  - Might apply data analytic techniques to evaluate alternative strategies

Modifying or replacing DIF method

- Alternative models worth further exploration
  - Recent contractor work considered wide range of alternatives
    - Better performing models could be refined
      - Account for known relationships among features
      - Account for actual threshold used to define significant audit adjustment
    - Pilot testing
- Need to become more explicit about importance placed on the 3 different performance metrics
  - Can better tune DIF or other methods to achieve desired performance
- New alternatives to DIF may be needed given trends in NRP design
  - May need to incorporate some operational audit data

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

Session 3. Developments in Technology and Analytics

Moderator:	***Terry Ashley*** *IRS: Taxpayer Advocate Service*
New Approaches to Estimating the Extent of Nonfiling	***Alan Plumley*** *IRS: RAAS*
Using Uplift Modeling to Improve ACS Case Selection and Compliance Outcomes	***Jan Millard*** *IRS: RAAS*
Recent IRS Discriminant Function (DIF) Model Improvements	***Getaneh Yismaw*** *IRS: RAAS*
Discussant:	***Brian Erard*** *Brian Erard & Associates*

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

Session 4. Enhancing Taxpayer Customer Experience

Moderator:	***Christine Oehlert*** *IRS: RAAS*
Increasing Take-up of the American Opportunity Tax Credit	***Anne Herlache*** *IRS: RAAS*
Customer Experience for Small Business and Self-Employed Taxpayers	***Kahoa Bonhomme*** *IRS: SB/SE*
Are Annual Federal Employment Tax Returns Effective? An Economic Analysis of Filing, Reporting, and Payment Compliance Associated with Forms 943 and 944	***Yan Sun*** *IRS: RAAS*
Discussant:	***Jacob Goldin*** *Stanford Law School*

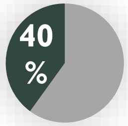
INCREASING TAKE-UP OF THE AMERICAN OPPORTUNITY TAX CREDIT

PRESENTED BY ANNE HERLACHE (ANNE.D.HERLACHE@IRS.GOV)
WITH CRYSTAL HALL & MARY CLAIR TURNER

DISCLAIMER: The views and opinions in this presentation reflect those of the authors. They do not necessarily reflect the views or the official position of the Internal Revenue Service or General Services Administration

BACKGROUND

- **The American Opportunity Tax Credit (AOTC)**
  - $2500 per student max, $1000 refundable



- **Barriers to take-up**
  - Awareness
  - Complexity
  - Coordination between dependent student & primary filer
  - Timing (calendar year vs academic year)

BACKGROUND

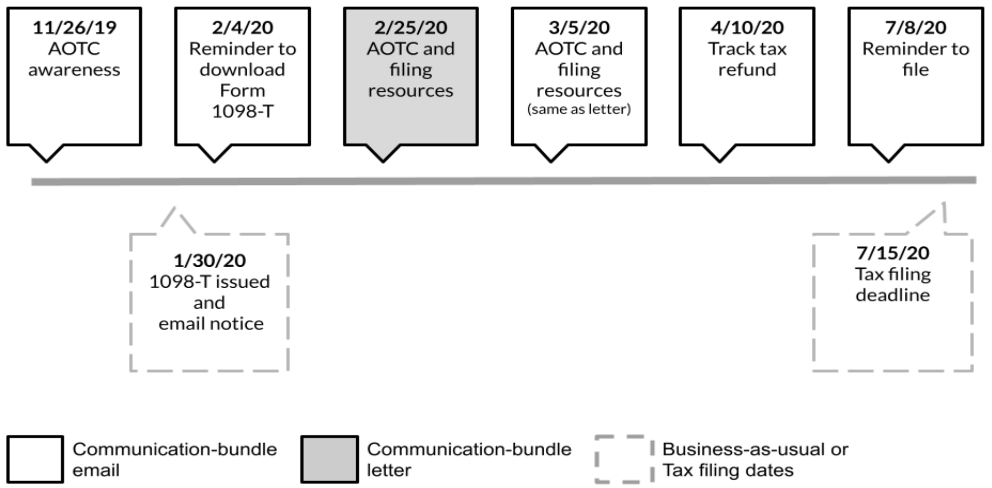
- **The American Opportunity Tax Credit (AOTC)**
  - $2500 per student max, $1000 refundable

- **Barriers to take-up**
  - Awareness
  - Complexity
  - Coordination between dependent student & primary filer
  - Timing (calendar year vs academic year)

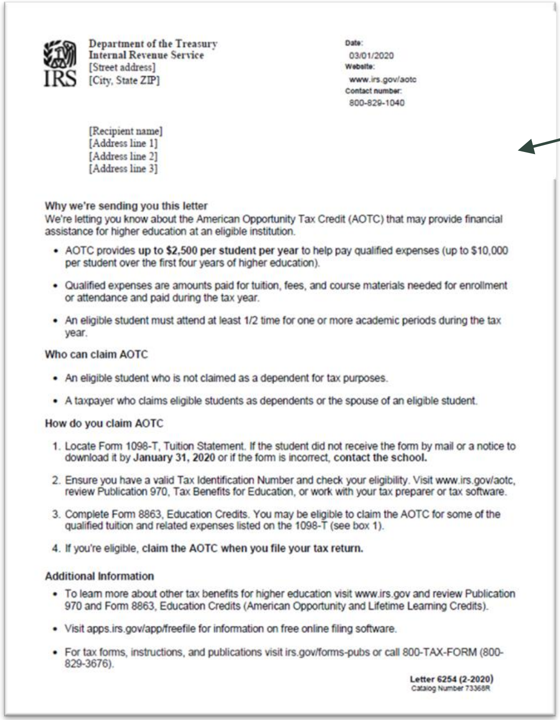
Coordinated Communication Bundle

DESIGN & TIMELINE

Treatment Assignment	Description
Business-as-usual (BAU)	University email: Form 1098-T is ready to download
Comm Bundle (IRS Letter)	BAU email Email bundle: 5 Emails to students (and authorized payer if applicable) IRS letter
Comm Bundle (University Letter)	BAU email Email bundle: 5 Emails to students (and authorized payer if applicable) University letter



COMMUNICATION BUNDLE EXAMPLES



Department of the Treasury
Internal Revenue Service
[Street address]
[City, State ZIP]

IRS

Date:
03/01/2020
Website:
www.irs.gov/aotc
Contact number:
800-829-1040

[Recipient name]
[Address line 1]
[Address line 2]
[Address line 3]

Why we're sending you this letter
We're letting you know about the American Opportunity Tax Credit (AOTC) that may provide financial assistance for higher education at an eligible institution.

- AOTC provides **up to $2,500 per student per year** to help pay qualified expenses (up to $10,000 per student over the first four years of higher education).
- Qualified expenses are amounts paid for tuition, fees, and course materials needed for enrollment or attendance and paid during the tax year.
- An eligible student must attend at least 1/2 time for one or more academic periods during the tax year.

Who can claim AOTC

- An eligible student who is not claimed as a dependent for tax purposes.
- A taxpayer who claims eligible students as dependents or the spouse of an eligible student.

How do you claim AOTC

1. Locate Form 1098-T, Tuition Statement. If the student did not receive the form by mail or a notice to download it by **January 31, 2020** or if the form is incorrect, **contact the school.**
2. Ensure you have a valid Tax Identification Number and check your eligibility. Visit www.irs.gov/aotc, review Publication 970, Tax Benefits for Education, or work with your tax preparer or tax software.
3. Complete Form 8863, Education Credits. You may be eligible to claim the AOTC for some of the qualified tuition and related expenses listed on the 1098-T (see box 1).
4. If you're eligible, **claim the AOTC when you file your tax return.**

Additional Information

- To learn more about other tax benefits for higher education visit www.irs.gov and review Publication 970 and Form 8863, Education Credits (American Opportunity and Lifetime Learning Credits).
- Visit apps.irs.gov/app/freefile for information on free online filing software.
- For tax forms, instructions, and publications visit irs.gov/forms-pubs or call 800-TAX-FORM (800-829-3676).

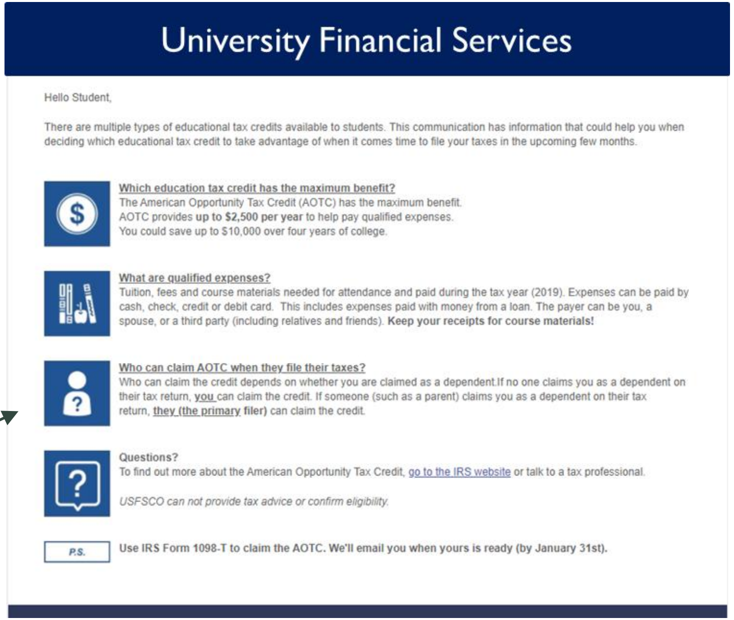
Letter 6254 (2-2020)
Catalog Number 73368R

Letters

- Streamlined--but thorough--information
- Sender varies (IRS or university)

Emails

- Personalized
- Clickable links
- Tested for viewing on different OS & devices



University Financial Services

Hello Student,

There are multiple types of educational tax credits available to students. This communication has information that could help you when deciding which educational tax credit to take advantage of when it comes time to file your taxes in the upcoming few months.

Which education tax credit has the maximum benefit?
The American Opportunity Tax Credit (AOTC) has the maximum benefit.
AOTC provides **up to $2,500 per year** to help pay qualified expenses.
You could save up to $10,000 over four years of college.

What are qualified expenses?
Tuition, fees and course materials needed for attendance and paid during the tax year (2019). Expenses can be paid by cash, check, credit or debit card. This includes expenses paid with money from a loan. The payer can be you, a spouse, or a third party (including relatives and friends). **Keep your receipts for course materials!**

Who can claim AOTC when they file their taxes?
Who can claim the credit depends on whether you are claimed as a dependent.If no one claims you as a dependent on their tax return, you can claim the credit. If someone (such as a parent) claims you as a dependent on their tax return, they (the primary filer) can claim the credit.

Questions?
To find out more about the American Opportunity Tax Credit, go to the IRS website or talk to a tax professional.

USFSCO can not provide tax advice or confirm eligibility.

P.S. **Use IRS Form 1098-T to claim the AOTC. We'll email you when yours is ready (by January 31st).**

SAMPLE

- Undergraduate students who are:
  - US citizens
  - Enrolled at least ½ time in one or more academic terms in CY 2019
  - Excluding:
    - Students with > 120 transfer credits
    - Students enrolled only in graduate programs
    - Students missing key information (e.g., SSN)

Treatment Assignment	Description	Sample Size
Business-as-usual (BAU)	University email: Form 1098-T is ready to download	9,530
Comm Bundle (IRS Letter)	BAU + Email Bundle + IRS Letter	4,776
Comm Bundle (University Letter)	BAU + Email Bundle + University Letter	4,765
Total communications-bundle sample		9,541
Total sample		19,071

ANALYSES

- **Research Question #1.** *Does sending a coordinated communications-bundle about AOTC to enrolled or recently enrolled college students increase take-up of AOTC?*

$$Y_{ib} = \beta_0 + \beta_1(T_{ib}) + \pi Z'_{ib} + \alpha_b + \varepsilon_{ib}$$

where *i* indexes students in blocks *b,* and

Y_{ib}: is the outcome of interest;

T_{ib} : is an indicator for random assignment to a communications-bundle group;

Z'_{ib}: is a vector of student-level characteristics;

α_{b}: are block fixed effects ; and

ε_{ib}: is an error term.

ANALYSES

- **Research Question #2.** *Are there messenger effects on AOTC take-up of being sent a coordinated communications-bundle with a letter from the IRS compared to being sent a coordinated communications-bundle with a letter from University?*

$$Y_{ib} = \beta_0 + \beta_1(IRS_{ib}) + \beta_2(University_{ib}) + \pi Z'_{ib} + \alpha_b + \varepsilon_{ib}$$

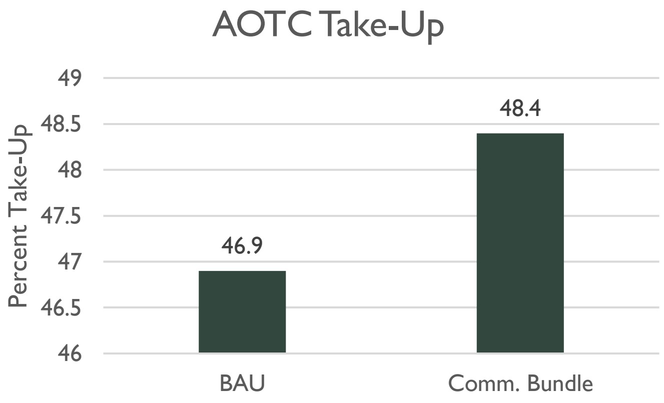
where *i* indexes students in blocks *b,* and

Y_{ib}:	is the outcome of interest;
IRS_{ib} :	is an indicator for random assignment to with communications-bundle group with IRS letter;
$University_{ib}$:	is an indicator for random assignment to with communications-bundle group with University letter;
Z'_{ib}:	is a vector of student-level characteristics;
α_{b}:	are block fixed effects for student enrollment
and	financial characteristics; and
ε_{ib}:	is an error term.

PRELIMINARY RESULTS

- **Research Question #1.** *Does sending a coordinated communications-bundle about AOTC to enrolled or recently enrolled college students increase take-up of AOTC?*

- The communication bundle significantly increases AOTC take-up by 1.5 pp

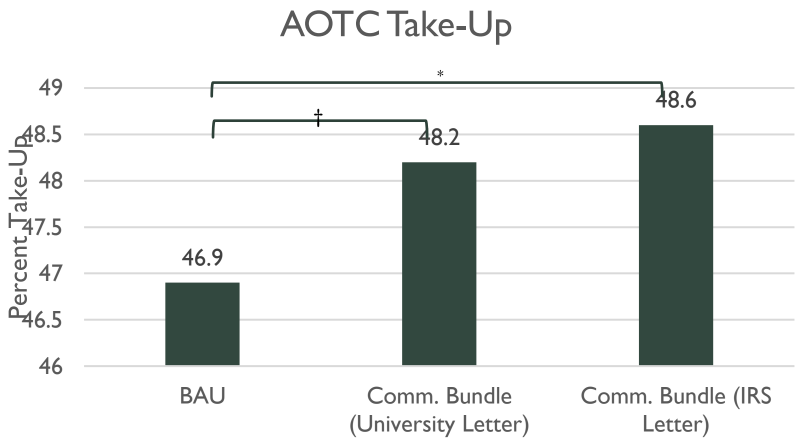


p = 0.02; 95% CI (0.2 - 2.8)

PRELIMINARY RESULTS

- **Research Question #2.** *Are there messenger effects on AOTC take-up of being sent a coordinated communications-bundle with a letter from the IRS compared to being sent a coordinated communications-bundle with a letter from University?*

- Communications bundle (IRS Letter) significantly increases AOTC take-up by 1.7 pp
- Communications bundle (University Letter) prompted a marginally significant increase in AOTC take-up by 1.3 pp
- There was no statistically significant difference in effects on AOTC take-up based on sender of letter

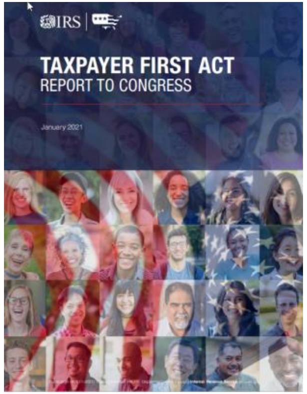


* BAU vs. Comm. Bundle w. IRS Letter p=0.04; 95% CI(0.1 - 3.3)

† BAU vs. Comm. Bundle w. University Letter p=0.09; 95% CI(-0.2 - 3.0)

NS: Comm. Bundle comparison IRS vs. University letter groups (p=0.75)

CONCLUSION



- Enhancing taxpayer customer experience
  - Empowering & enabling taxpayers
  - Timely communication
- Informative for the IRS & postsecondary institutions

- Follow-up analyses with the current project
  - Preparation method
  - Filing status
  - Timing of filing
  - Email engagement metrics
  - Educational outcomes

Thank You

Contact: Anne.D.Herlache@IRS.gov

The Customer Experience for Small Business and Self-Employed Taxpayers

Kahoa Bonhomme

Internal Revenue Service - SB/SE Research

IRS-TPC Joint Research Conference | June 2021

Topics

- Overview
- The SB/SE Taxpayer
- SB/SE Taxpayer Subpopulations
- Future Research

Overview

SB/SE Research seeks to integrate its actionable research efforts to achieve a better understanding of the future customer experience and to inform a more inclusive customer experience framework.

Objective

To gain a **DEEPER UNDERSTANDING** of the SB/SE taxpayer experience

Research Integration

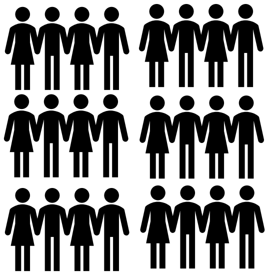
- Focus Groups
- Taxpayer Phone Call Analysis
- Surveys
- Journey Maps

Desired Outcome

Identification of **IMPROVEMENT OPPORTUNTIES** for the SB/SE taxpayer experience

The SB/SE Taxpayer Population

57
million
SB/SE
taxpayers

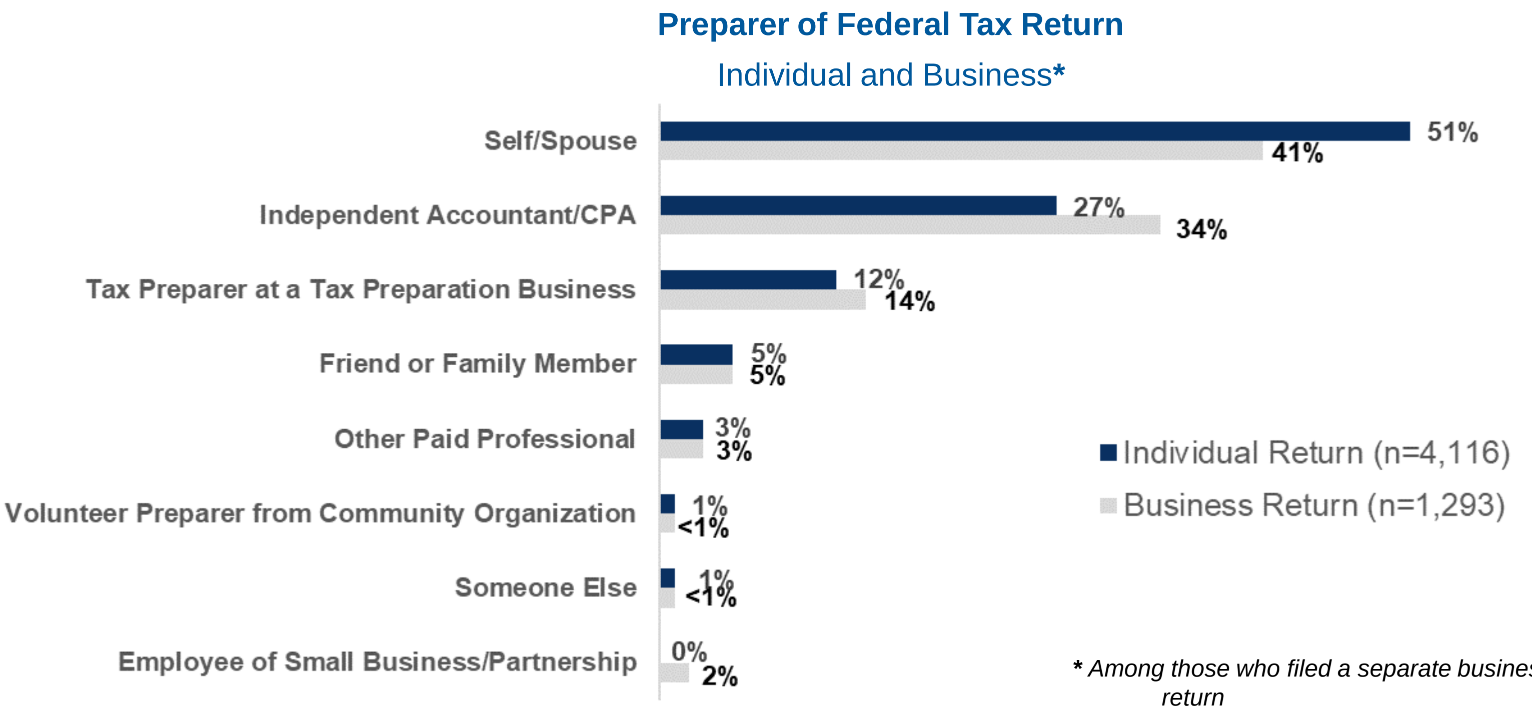


SB/SE Segment	**Size** (# of Taxpayers)
Schedule C, E, or F filers	**47** million
Corporation filers	**6.8** million
Partnership filers	**3.8** million
Employment Tax return filers	**26.8** million
Excise Tax return filers	**1.1** million
Gift Tax return filers	**250,000**
Estate Tax return filers	**27,000**

1/3
of the overall
taxpayer
base

The SB/SE Taxpayer: Preparation of Tax Returns

Business tax returns are less likely to be self-prepared than individual returns.

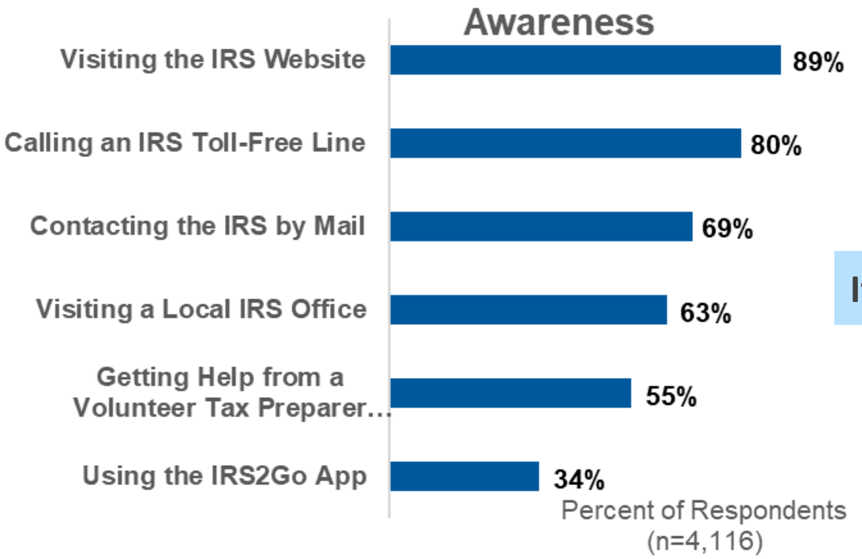


** Among those who filed a separate business return*

Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

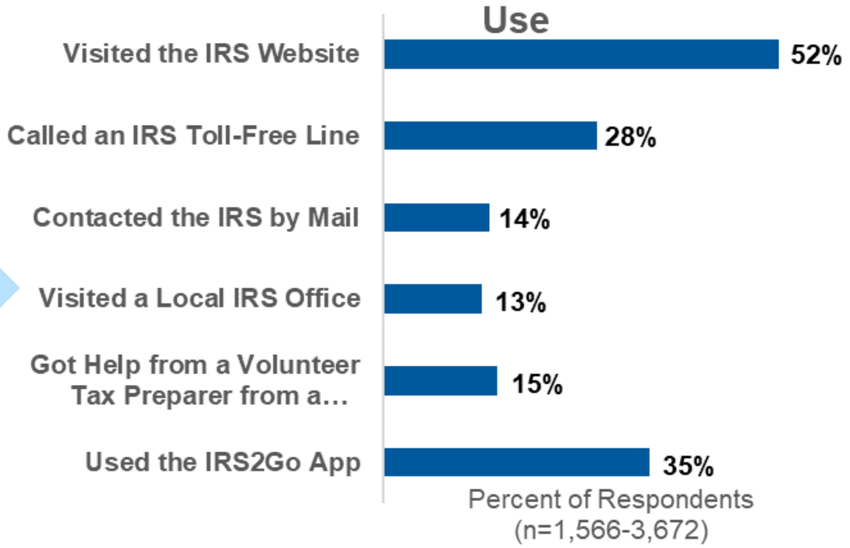
Overall awareness of IRS service channels is high.

Awareness of IRS Channels



If aware

Use of IRS Channels [among those aware]



Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

SB/SE Taxpayer Subpopulations: Overview

In addition to the overall SB/SE population, five subpopulations were studied.

SB/SE Subpopulations

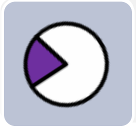
[All percentages are from the 2020 *Customer Experience, Expectations, and Needs Survey* National Report.]

17% of SB/SE respondents were **Balance Due Taxpayers**.

25% of SB/SE respondents paid employees or independent contractors: **Employment Tax Filers**.

15% of SB/SE respondents were **Gig Workers**.

9% of SB/SE respondents were **Spanish-Preferred Taxpayers**.



21% of SB/SE respondents have been in business for less than 1 year: **New SB/SE Taxpayers.**

SB/SE Balance Due Taxpayers

For many taxpayers, IRS.gov is the preferred IRS service channel.

Subpopulation Definition

Taxpayers with a balance due that was not paid at the time of filing

CEEN Survey

IRS channel used for getting tax information

- IRS.gov: 93% of respondents
- IRS telephone line: 9% of respondents
- Local IRS office: 5% of respondents

Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

Taxpayer Telephone Calls

Many taxpayers called the IRS telephone line, after **first trying to resolve the issue online.**

Source: IRS Balance Due Taxpayer: Taxpayer Vocabulary, 2020.

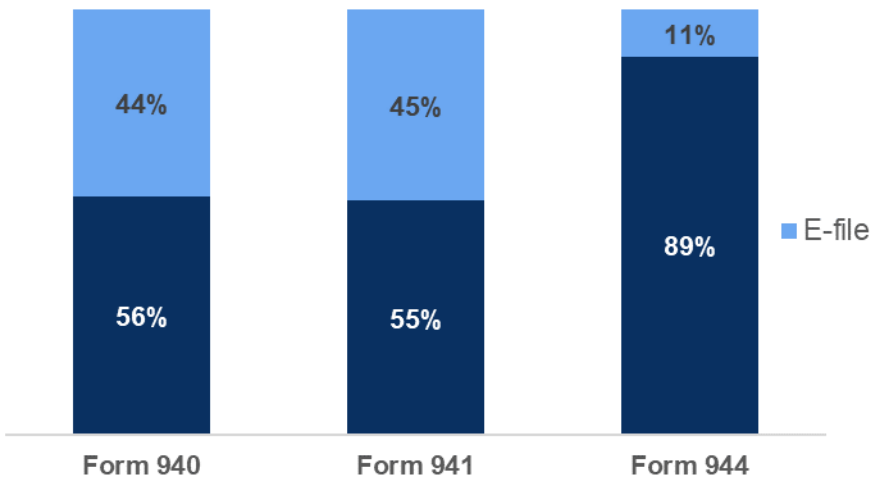
Preference: **Completing tasks online**

Employment tax returns have the lowest e-file rates among business returns.

Subpopulation Definition

Taxpayers who file employment tax returns

Employment Tax Returns: Filing Method
[Calendar Year 2018]



The **external factors** that impact e-filing are the perceived **ease of paper filing** and **payroll software/paid preparers encouraging paper filing.**

Sources:
RAAS Publication 6149 2019 Final: Calendar Year Return Projections by State 2019-2026, Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021, and Employment Tax: Environmental Scan, April 2021.

Journey maps and survey data reveal common pain points experienced by SB/SE employment tax filers.

Searching for tax information

Taxpayers may face challenges in navigating IRS and non-IRS resources.

“Accounting help or someone who can help figure out payments, forms or helping with questions”

“A better way to search for the information you need and have it be in more understandable language."

"It would be good if there was a better way to be able to see your IRS account with amounts due and payments.”

Making estimated tax payments

Taxpayers may find it difficult to make on-time payments and track payment status/history.

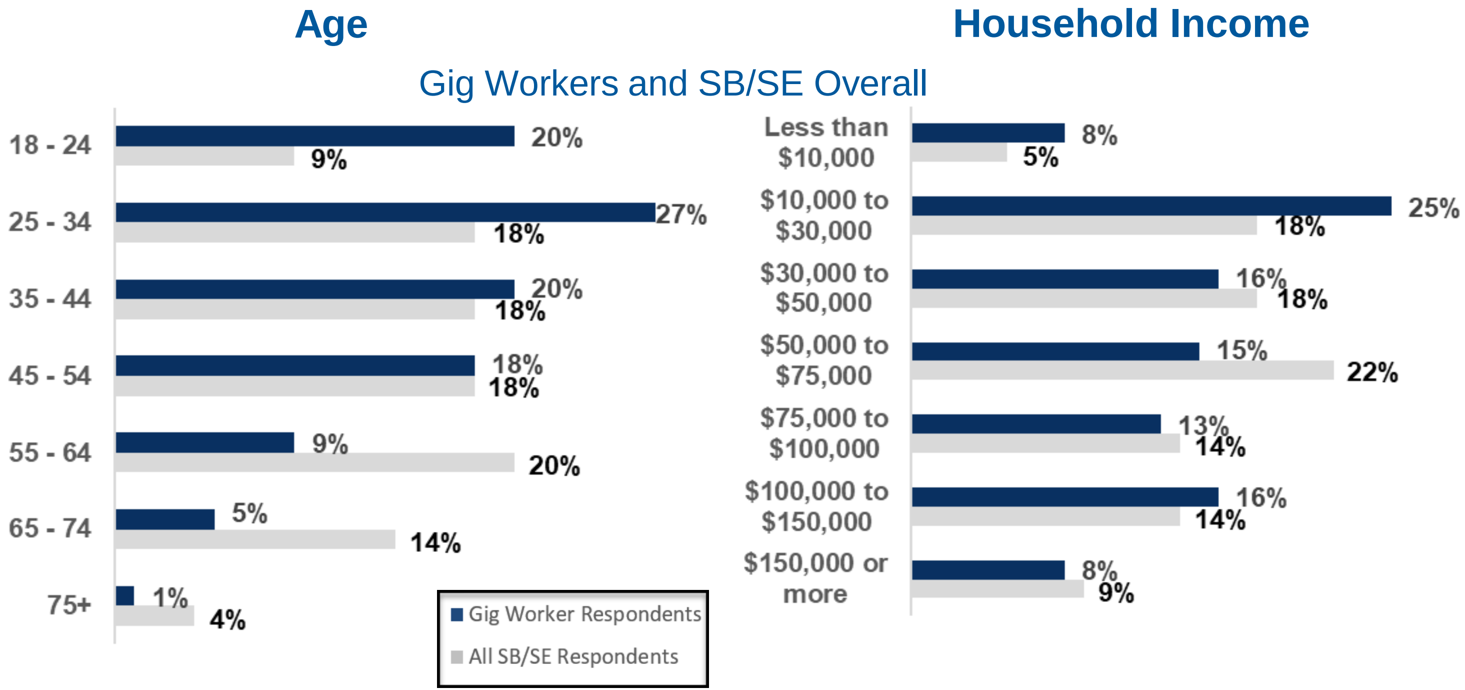
Source: Employment Tax Journey Map, 2020 and *Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021*

Gig workers differ from overall SB/SE taxpayers in key demographics.

Subpopulation Definition

Taxpayers who have income from gig work

Examples
rideshare driver, pet sitter, rental host



Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

SB/SE Gig Workers

Focus group discussions and journey mapping uncovered important aspects of the gig worker experience.

Instead of using IRS resources, gig workers generally turn to **Google, social media, tax professionals, and family/friends** for answers to their questions.

One of the biggest challenges gig workers face is identifying what is considered **taxable income.**

Gig workers need more guidance on what qualifies as an **acceptable deduction.**

Gig workers want more advice on **record-keeping.**

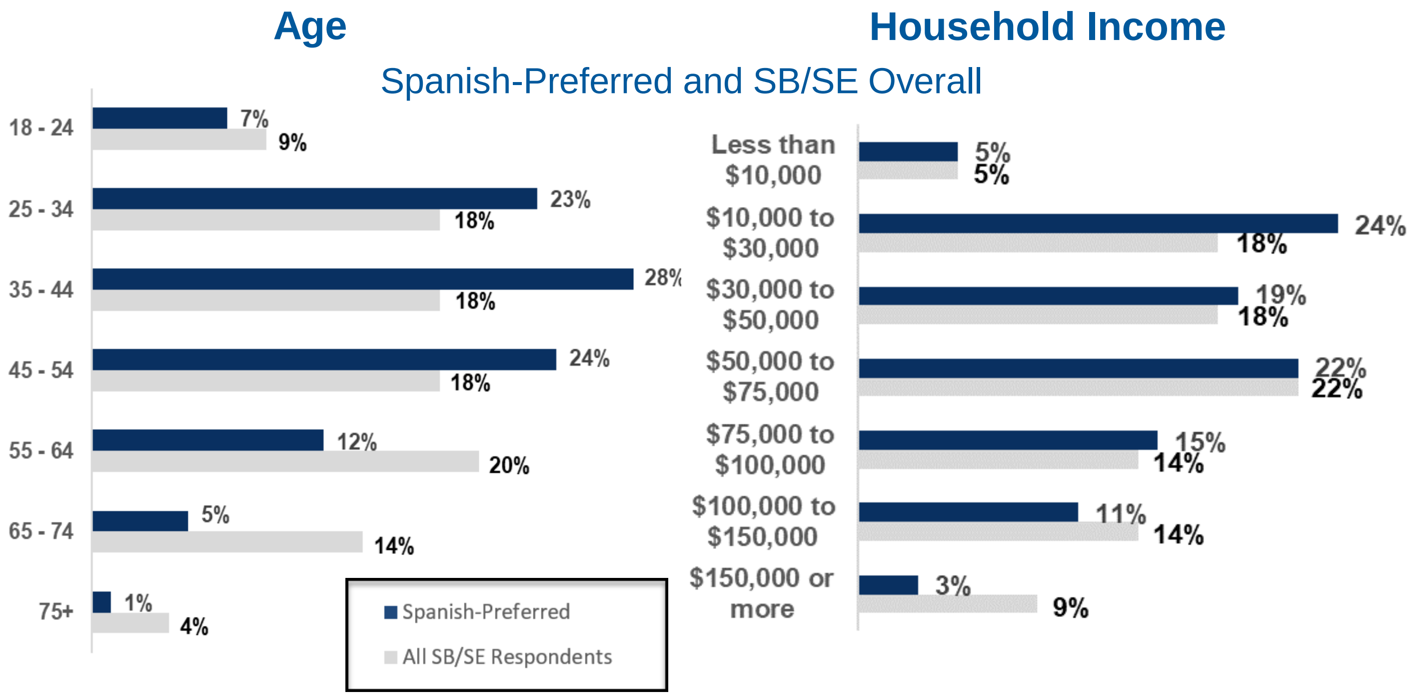
Source: Employment Tax Journey Map, 2020 and *Gig Economy Worker Focus Groups, 2021.*

SB/SE Spanish-Preferred Taxpayers

Spanish-preferred differs from overall SB/SE taxpayers in key demographics.

Subpopulation Definition

CEEN survey respondents who chose to take the survey in Spanish



Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

SB/SE Spanish-Preferred Taxpayers

In addition to demographic differences, there were differences between Spanish-preferred and overall SB/SE respondents in IRS-related issues.

Tax Preparation

Used a paid preparer for the TY2019 federal individual income tax return

63% *Spanish-Preferred*

42% *All SB/SE*

IRS Toll-Free Line

Used the IRS toll-free telephone line (among those aware of this channel)

47% *Spanish-Preferred*

28% *All SB/SE*

IRS Notices

Have ever received an IRS notice

43% *Spanish-Preferred*

31% *All SB/SE*

Outreach

Want the IRS to conduct outreach about tax filing requirements

66% *Spanish-Preferred*

28% *All SB/SE*

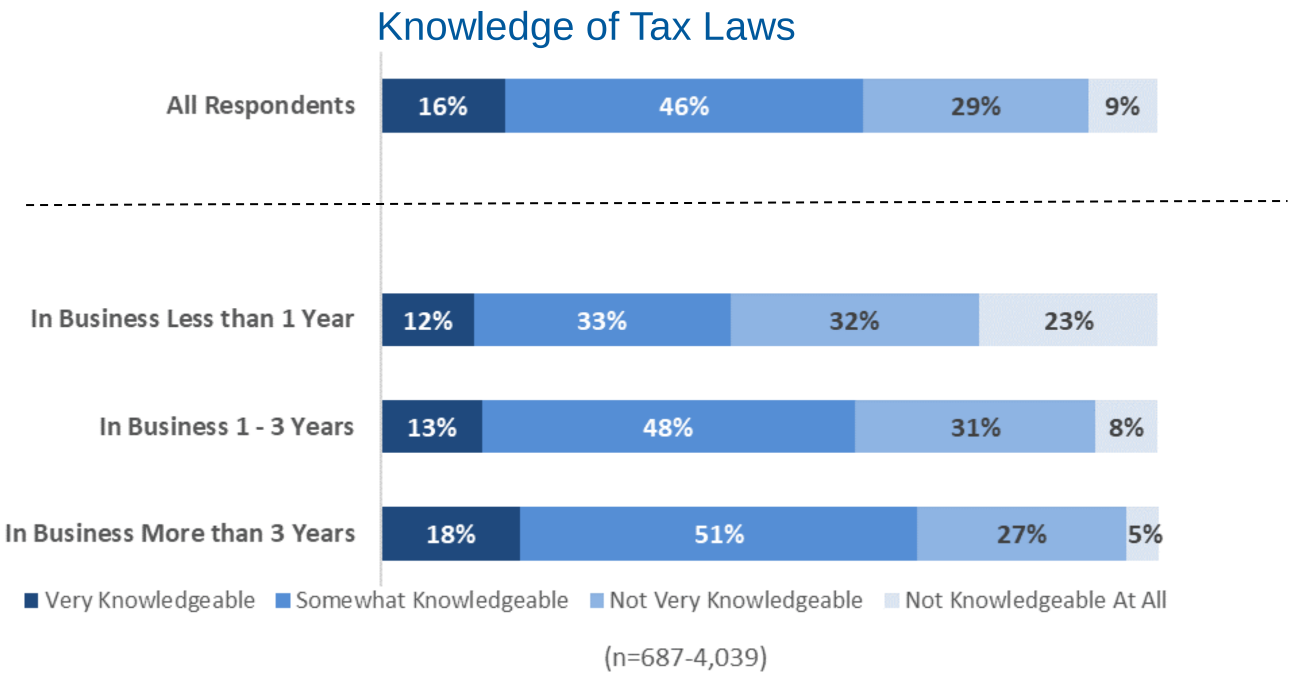
Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

New SB/SE Taxpayers

Self-reported knowledge of tax laws is lower for new SB/SE taxpayers.

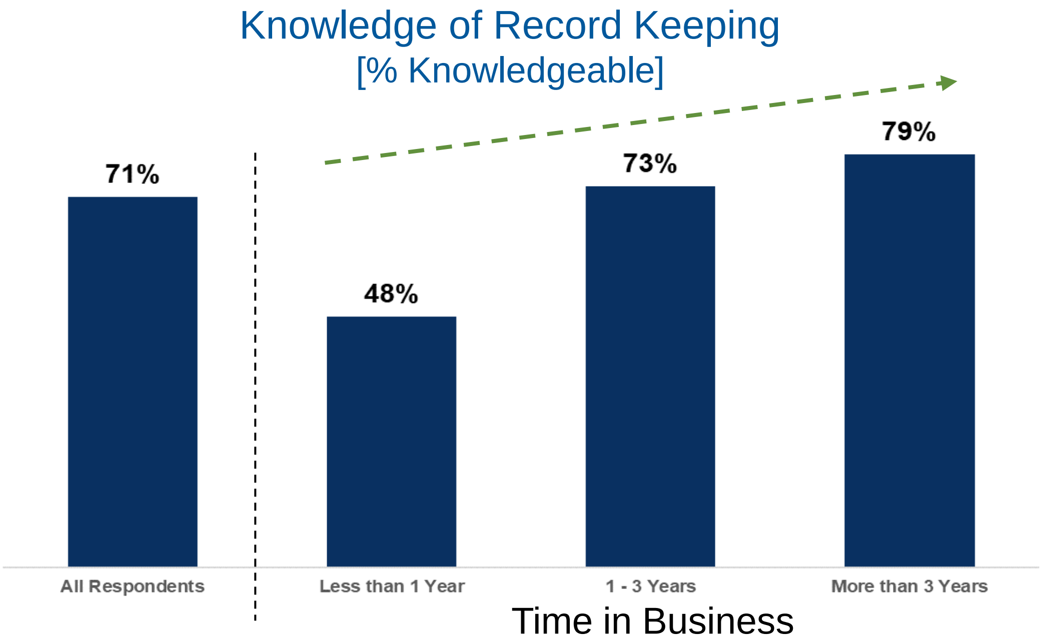
Subpopulation Definition

Taxpayers who have been a small business owner/partner or self-employed for less than one year



Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

Self-reported knowledge of record keeping is much lower for new SB/SE taxpayers than for more experienced small businesses.



"I wish I knew how important it was to keep receipts, keep your paperwork, and keep everything filed."

CEEN focus group participant

Source: Customer Experience, Expectations, and Needs Survey 2020 National Report, February 2021.

Future Research

Taxpayer First Act (TFA)

- The TFA was passed by Congress in 2019.
- The TFA's overarching objective is to improve IRS operations.
- A major component of the TFA is the requirement that the IRS develop a comprehensive customer service strategy.

How do the TFA provisions impact the SB/SE taxpayer experience?

Additional SB/SE Subpopulations

- Our review of multiple research projects uncovered important SB/SE subpopulations.
- Identifying and studying subpopulations within SB/SE can enhance outreach and other improvement efforts.

What other subgroups exist within SB/SE?

Are Annual Federal Employment Tax Returns Effective? An Economic Analysis of Filing, Reporting and Payment Compliance Associated with Forms 943 and 944

Yan Sun and Stephanie D. Needham,

with Contribution from Barbara Sandstrom, Mel Hadley, Terry Manzi, Lindsay Schrock, and Tiffany Zerangue

11th Annual IRS-TPC joint Research Conference

June 24, 2021

The Executive Champion:

Holly Donnelly (RAAS, IRS)
Imelda Deniz-Vazquez (LB&I, IRS)

The Innovation Lab Team Members:

Mel Hadley (SB/SE, IRS)
Stephanie D. Needham (RAAS, IRS)
Terry Manzi (RAAS, IRS)
Barbara Sandstrom (SB/SE, IRS)
Lindsay Schrock (RAAS, IRS)
Yan Sun (RAAS, IRS)
Tiffany Zerangue (RAAS, IRS)

Additional Acknowledgements:

Specialty Exam Policy, SB/SE, IRS
Customer Account Services (CAS), WB&I, IRS

Context

Employers are required to

- **Withhold federal income tax, and social security and Medicare taxes**
- **Pay and report these tax liabilities to IRS**
- **File the Form 941**
  - **If the employer pays wages subject to income tax withholding or social security and Medicare taxes, the entity must file Form 941 quarterly**

IRS provides certain groups an alternative filing option (i.e. annual)

- **Form 943**
  - **Farm employers who employ agricultural workers are permitted to file the annual Form 943 - Employer's Annual Federal Tax Return for Agricultural Employees (implemented in 1956)**
- **Form 944**
  - **Small employers with an annual tax liability less than $1,000 are permitted to file the annual Form 944 - Employer's Annual Federal Tax Return (introduced in 2006)**

Research Questions

- **Has the burden of the annual filing outweighed its benefits?**

Strategy

- **Understand the Forms 943 & 944 Filing Landscape: Questions Considered.**
  - ✓ Are employers filing the correct form?
  - ✓ Are IRS filing rules followed?
  - ✓ Is the employment tax filed timely?

- **Forms 943 & 944 Recommendation(s): Questions Considered.**
  - ✓ What are the benefits and burdens of filing, and processing Forms 943 & 944?
  - ✓ Are the annual forms still needed by the employers?
  - ✓ Do the annual forms need any modifications? And how?

Research Goals

- **Assess the effectiveness of the annual employment tax forms 943 & 944**
- **Provide data-supported evidence for potential regulatory changes**

Research Approaches and Data

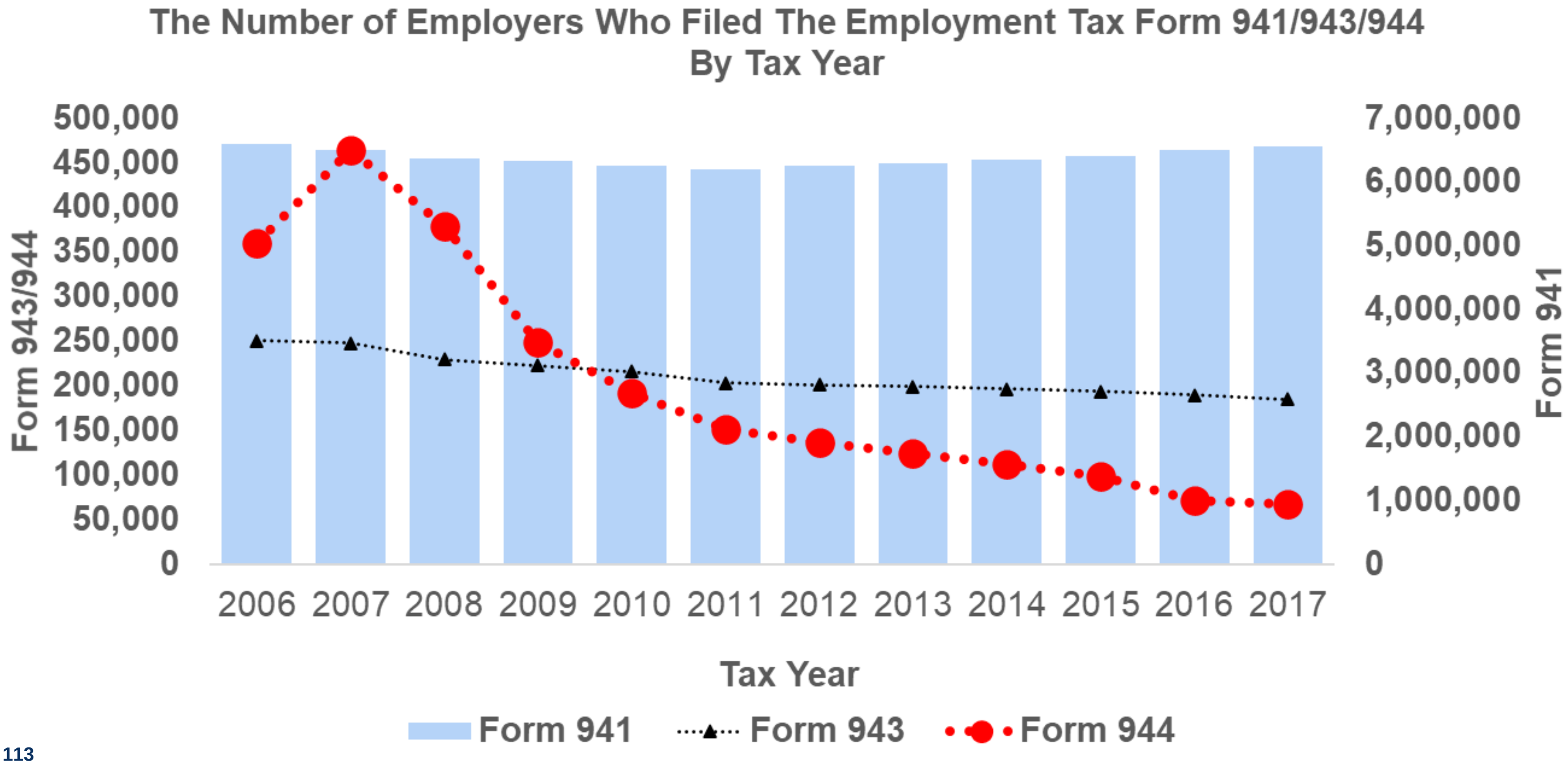
Research Approaches

- Use a comparative analysis approach to understand the Filing, Reporting and Payment Compliance of Forms 943 & 944
- Provide sufficient statistics to characterize the annual filing population
- Develop an empirical model to quantify the tax payment compliance

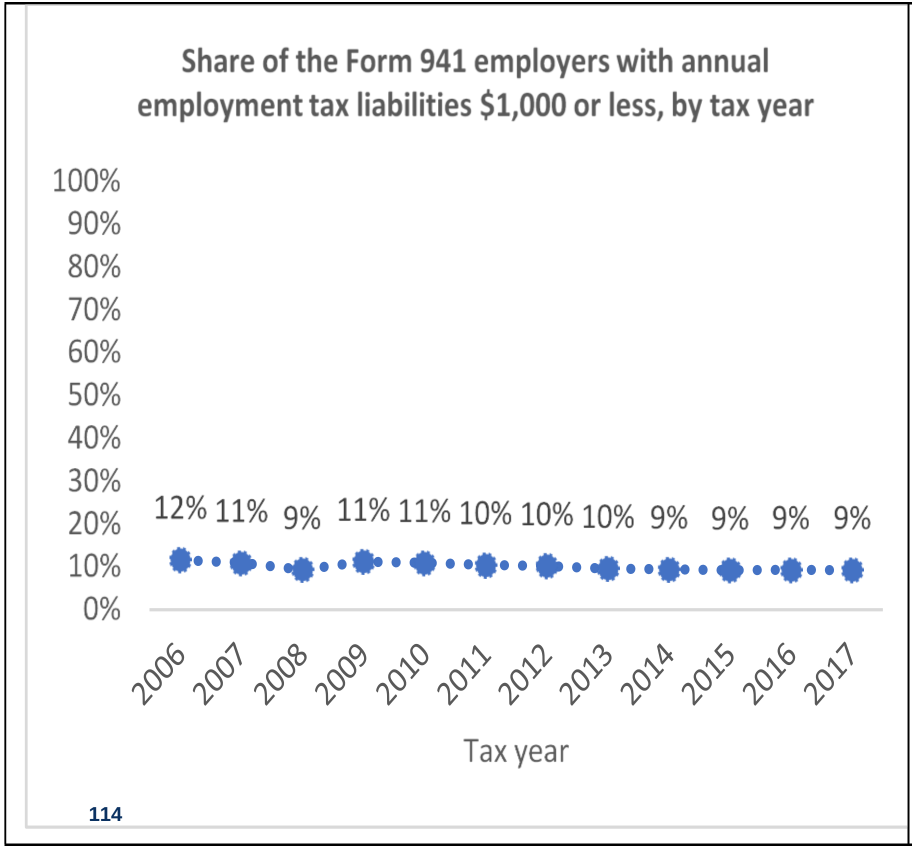
Data

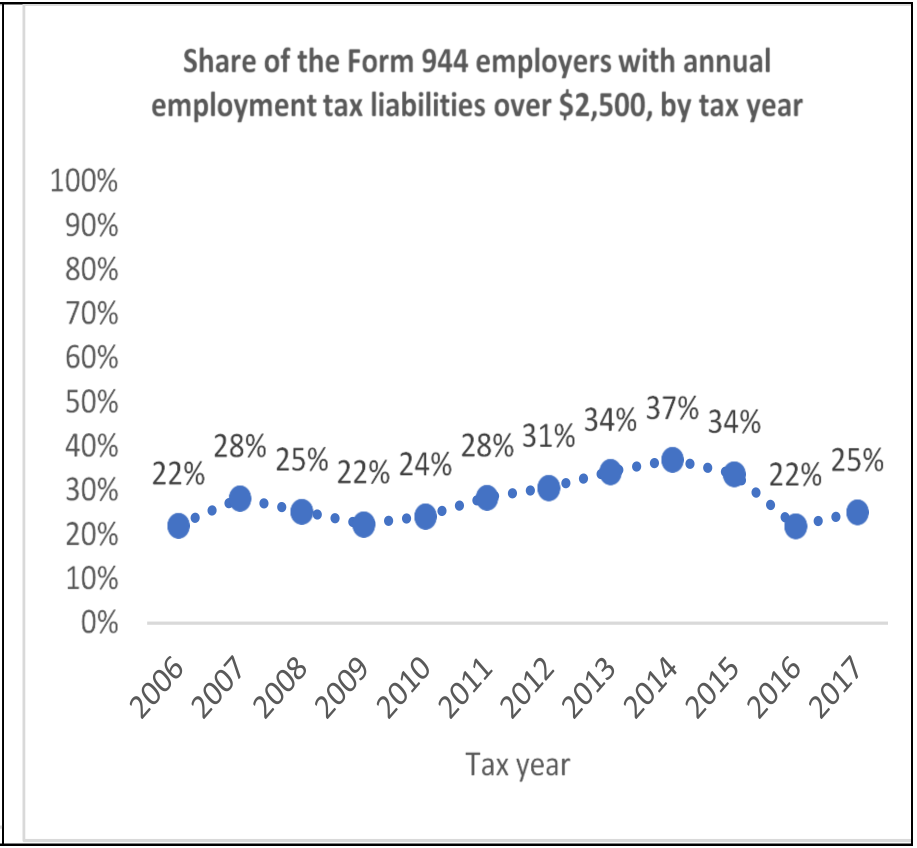
- The Compliance Data Warehouse (CDW) data
  - BRTF (Business Return Transaction File)
  - BMF (Business Master File)

Form 944 Filings Have Been in Steady Decline Over the Past Decade



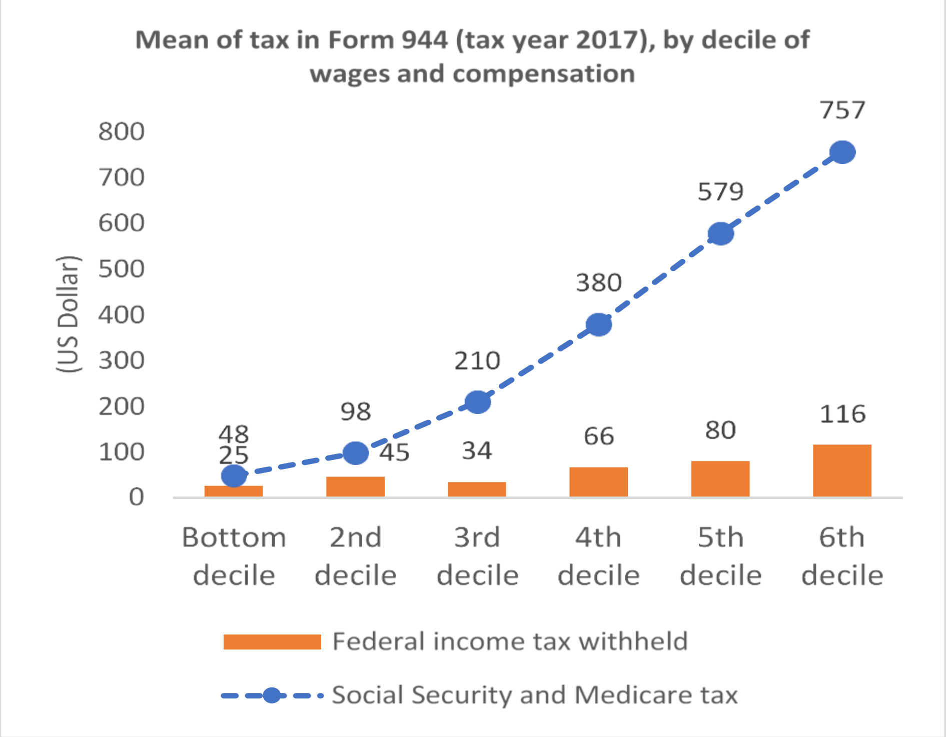
Majority of Small Businesses Filed Form 941, And A Non-Trivial Number of Large Businesses Filed Form 944



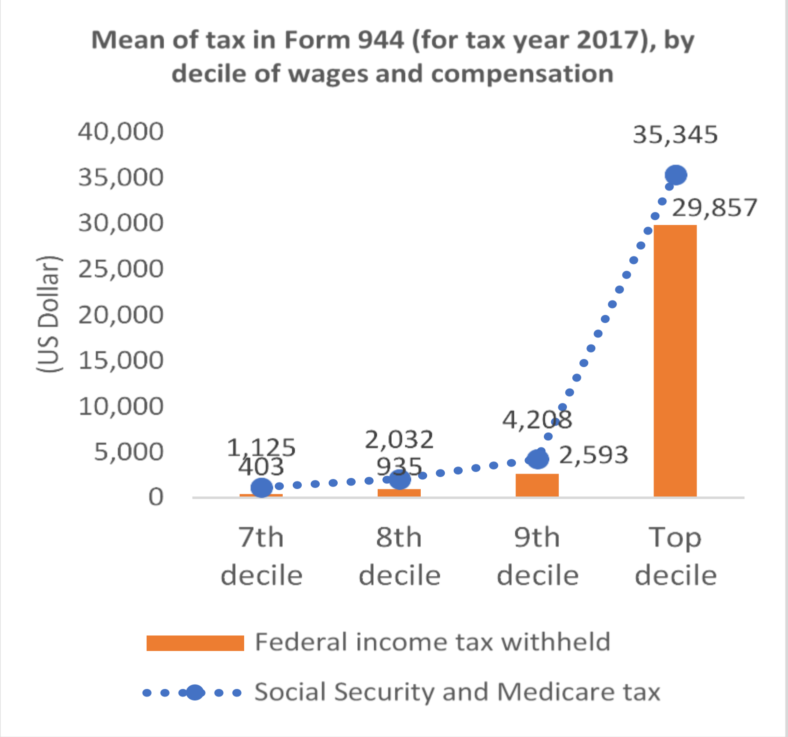


The Top Decile of the Form 944 Filing Population had a Significantly Higher Tax Liability Than the Bottom Deciles

Panel A: Bottom to 6th decile

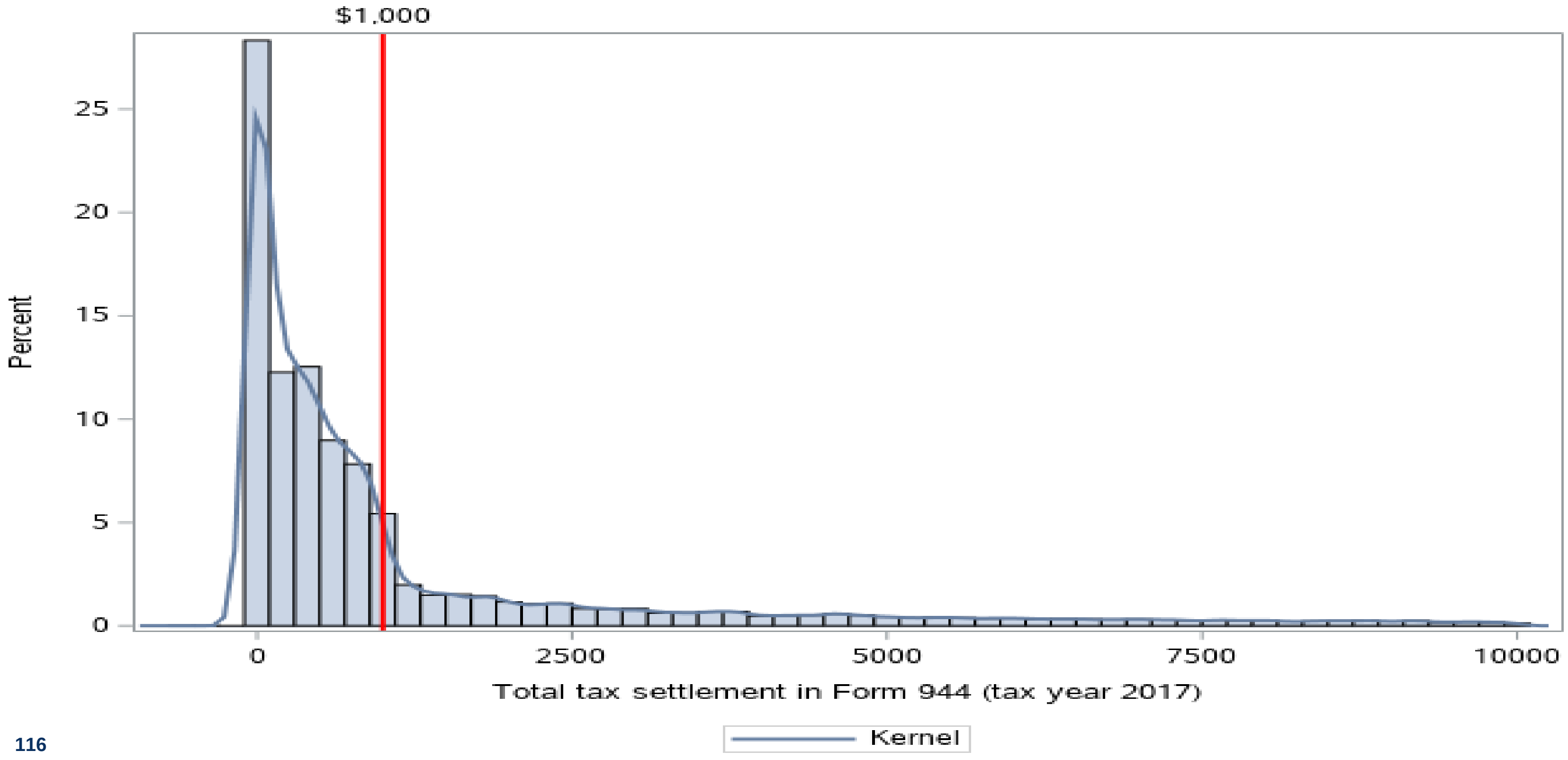


Panel B: 7th to top decile



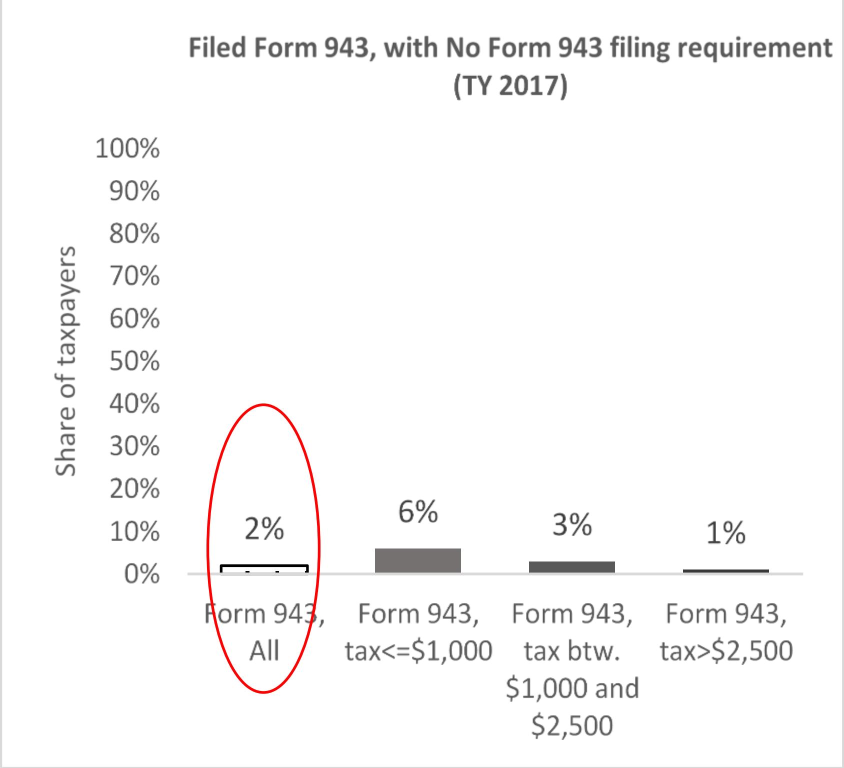
Approximately 12 Percent of Form 944 Filers in Tax Year 2017 Had An Annual Tax Liability Over $10,000

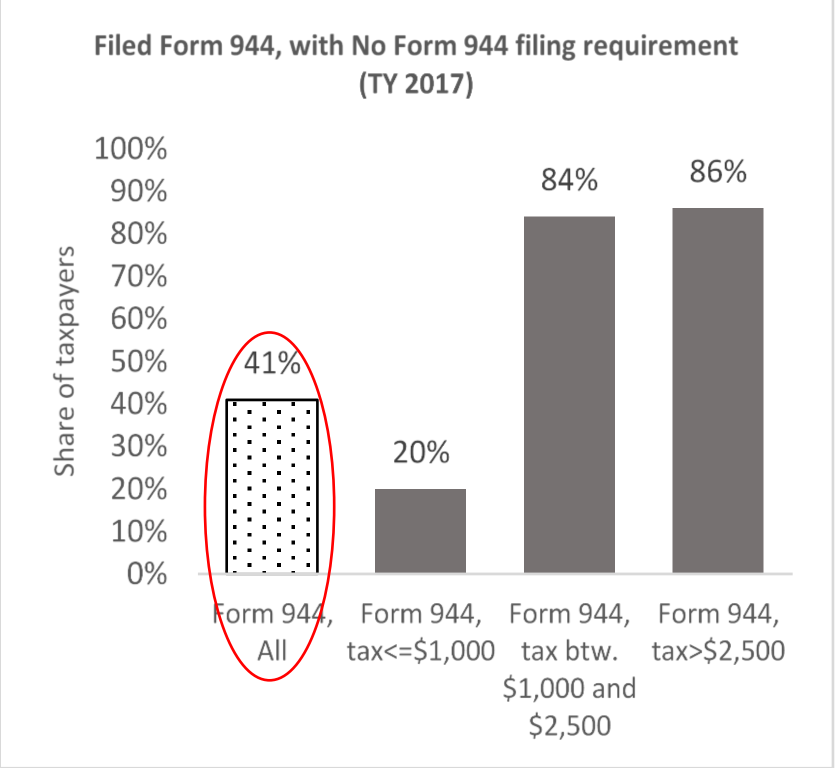
Truncated Distribution of Tax Liability of Form 944 Filing, Tax Year 2017



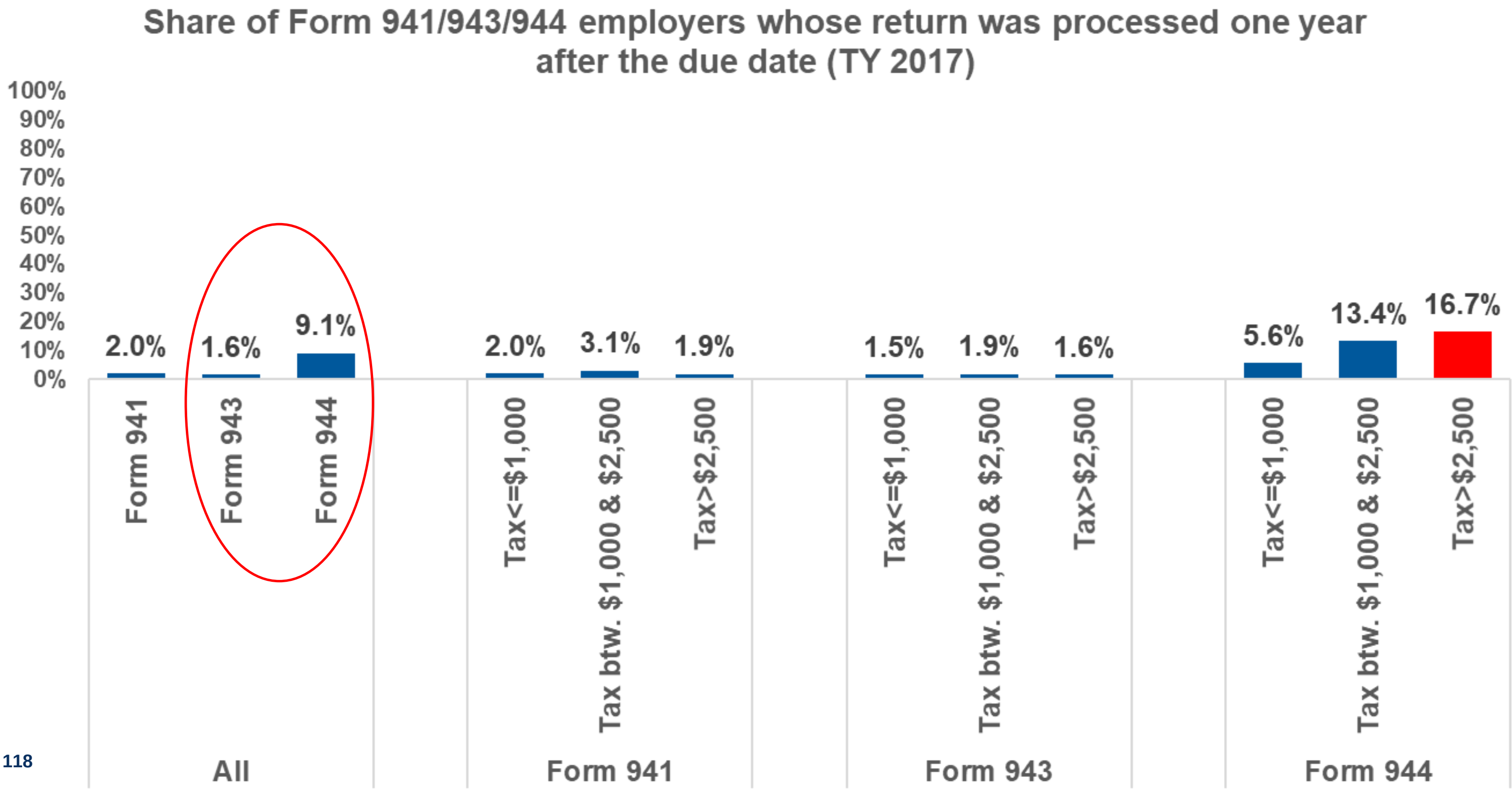
Form 943 Filers Were More Likely to Follow the IRS Tax Filing Requirement

Form 943	Form 944

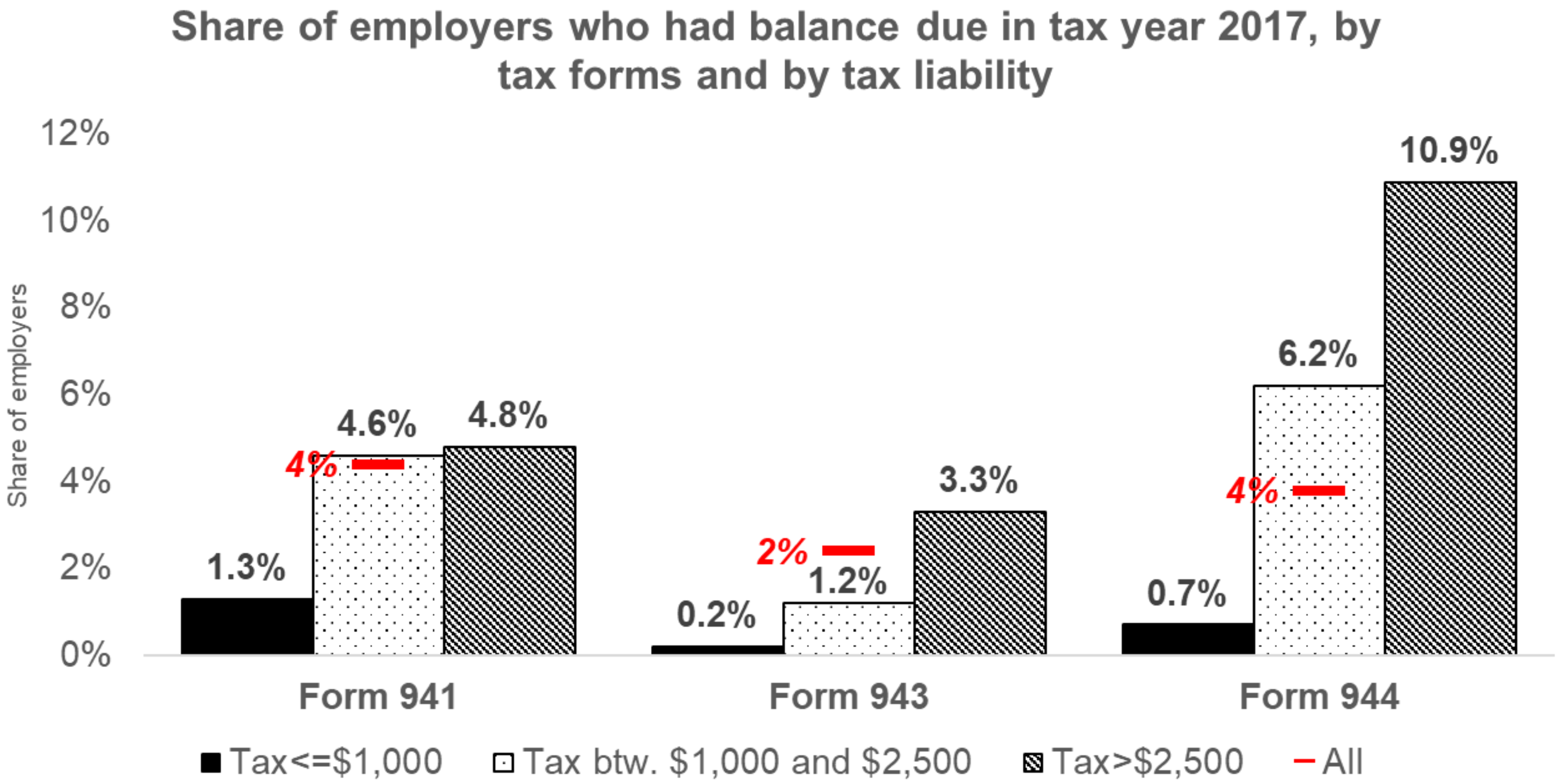




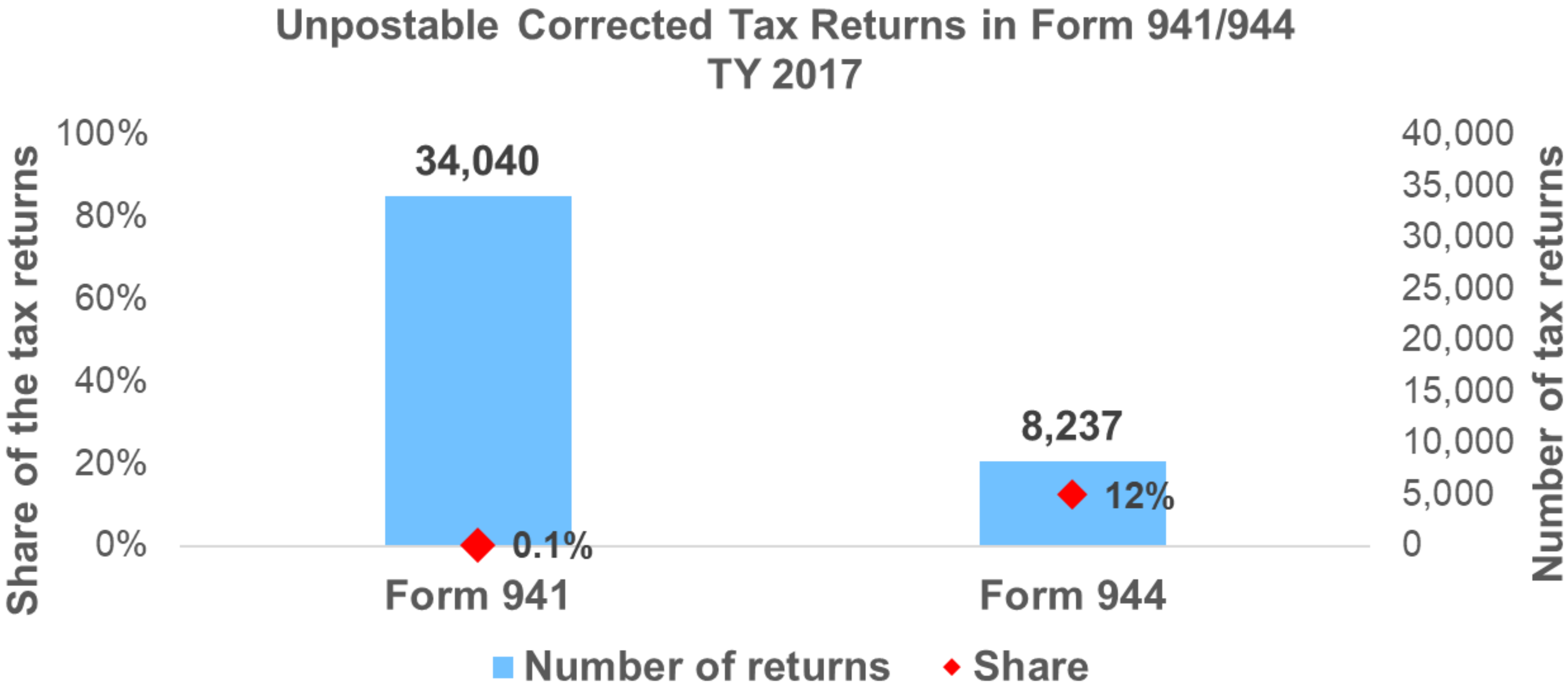
Form 944 Filers With Higher Tax Liability Were Less Compliant in Terms of Timely Filing



Form 944 Filers With Tax Liability Over $2,500 had the Highest Rate of Payment Delinquency



Form 944 Filers had a Significantly Higher Rate Of Unpostable Corrected Tax Returns



Notes

1. The "Unpostable corrected tax return" in this study refers to the unpostable returns, corrected, posted and captured by CDW
2. The CDW data does not show any Form 943 unpostable in Tax Year 2017.

Empirical Model: Logistic Regression

$$ln\frac{P_i^f}{1-P_i^f} = \beta_0^f + \sum_{j=1}^{k}\beta_j^f x_{ij}^f \qquad (1)$$

$$f = Form\ 943, Form\ 944$$

P_i^f = The probability of tax noncompliance of employer i for Form f

$$P_i^f\left(Y_i^f = 1 \middle| X_i^f = x_i^f\right) = \frac{e^{\beta_0^f+\sum_{j=1}^{k}\beta_j^f x_{ij}^f}}{1+e^{\beta_0^f+\sum_{j=1}^{k}\beta_j^f x_{ij}^f}} \qquad (2)$$

Y_i^f = A dichotomous outcome variable

Empirical Result

Explanatory variable	Parameter estimate	Standard error	Wald Chi-square	Pr>Chi-Sqr	Odds ratio
Panel A: Form 943					
Intercept	-7.8071	0.4356	321.19***	0.0001	
Log of total wages	0.1875	0.0132	200.93***	0.0001	1.21
Dummy, Tax liability $1,000 or under (base)					
Dummy, Tax liability btw. $1,000 and $2,500	0.6867	0.1525	20.29***	0.0001	1.99
Dummy, Tax liability above $2,500	1.2580	0.1398	80.97***	0.0001	3.51
Industry sectors fixed effects			Yes		
Geographical location fixed effects			Yes		
Observations			184,737		
Chi-Square Statistics of overall significance of the regression = 2305.35 (p-value=0.0001)					
Panel B: Form 944					
Intercept	-5.6035	0.3760	222.06***	0.0001	
Log of total wages	0.1961	0.0207	89.63***	0.0001	1.22
Dummy, Tax liability $1,000 or under (base)					
Dummy, Tax liability btw. $1,000 and $2,500	1.6441	0.0916	322.19***	0.0001	5.17
Dummy, Tax liability above $2,500	1.9289	0.0943	418.72***	0.0001	6.88
Industry sectors fixed effects			Yes		
Geographical location fixed effects			Yes		
Observations			66897		
Chi-Square Statistics of overall significance of the regression = 2199.24 (p-value=0.0001)					

Note: *** Statistically significance at p-value<0.01

Model Interpretation – Form 943 & 944

- The implementation and processing of Form 944, whose eligibility is based on the employer's annual tax liability, is more complex than Form 943, whose eligibility is based on the industry sector

- Form 943 & 944 payment delinquency is positively correlated with the employer's annual tax liability

- Large businesses that file Form 944 are approximately twice as likely to have tax due noncompliance than large businesses that file Form 943

Findings

- Confirmed problems identified by IRS Subject Matter Experts (SMEs) in SB/SE
- Observed divergent trends for the filing populations of the annual forms
- Identified that the Form 944 has not been widely adopted by small businesses since its introduction in 2006
- Found a non-trivial number of large businesses that filed the Form 944
- Demonstrated that Form 943 filers exhibit less filing confusion than Form 944 filers
- Showed that Form 944 filers with larger tax liabilities have higher rates of tax noncompliance

Recommendations

Form 944 Recommendation: Eliminate Form 944.

Form 943 Recommendation : No Form changes, or additional action recommended.

Employment tax filing in Tax Year 2017, by tax form and by tax liability

Employment tax	Form 941		Form 943		Form 944	
liability	No. of employers	Share	No. of employers	Share	No. of employe	Share
All	6,562,139	100%	184,737	100%	66,897	100%
Tax<=$1,000	605,407	9%	39,651	21%	43,786	65%
Tax btw. $1,000 & $2,500	406,493	6%	21,878	12%	6,513	10%
Tax>$2,500	5,550,239	85%	123,208	67%	16,598	25%

Discussion of Papers in Session 4

Jacob Goldin, Stanford Law School

June 24, 2021

Paper 1

Hall, Herlache, and Turner: AOTC Take-Up

This paper has many strengths

- Strong methodology
- Unique collaboration
- Multiple interesting questions

Broader Context?

- What is the scope of the problem?
  - National take-up rate (share of individuals receiving 1098-T who claim)
  - Can’t observe all eligibility factors (drug conviction) but pretty close
  - Do demographic correlates of take-up match heterogeneity in estimated treatment effect

- How does the experimental sample match the broader population?
  - Demographics
  - Take-up rates

Role of Tax Filing?

- How much of incomplete take-up is driven by non-filing?
- Why it matters:
  - Is it important to raise awareness of the credit or is increasing filing enough to raise take-up?
- Policy interventions differ depending on role of non-filing
  - Simplified claiming process, like CP-27 for EITC (confirm non-felon status)

- Related: If interventions increases filing, may be spillovers to other credits

Other Comments

- Why not look at variation in whether sent to student or parent?
- Limit to apparently eligible based on information returns?
- Assess mechanism by looking at subgroups? E.g., probably not unawareness for people who claimed in past

Paper 2

Bonhomme, Holland, and Hu:
SB/SE Customer Experience

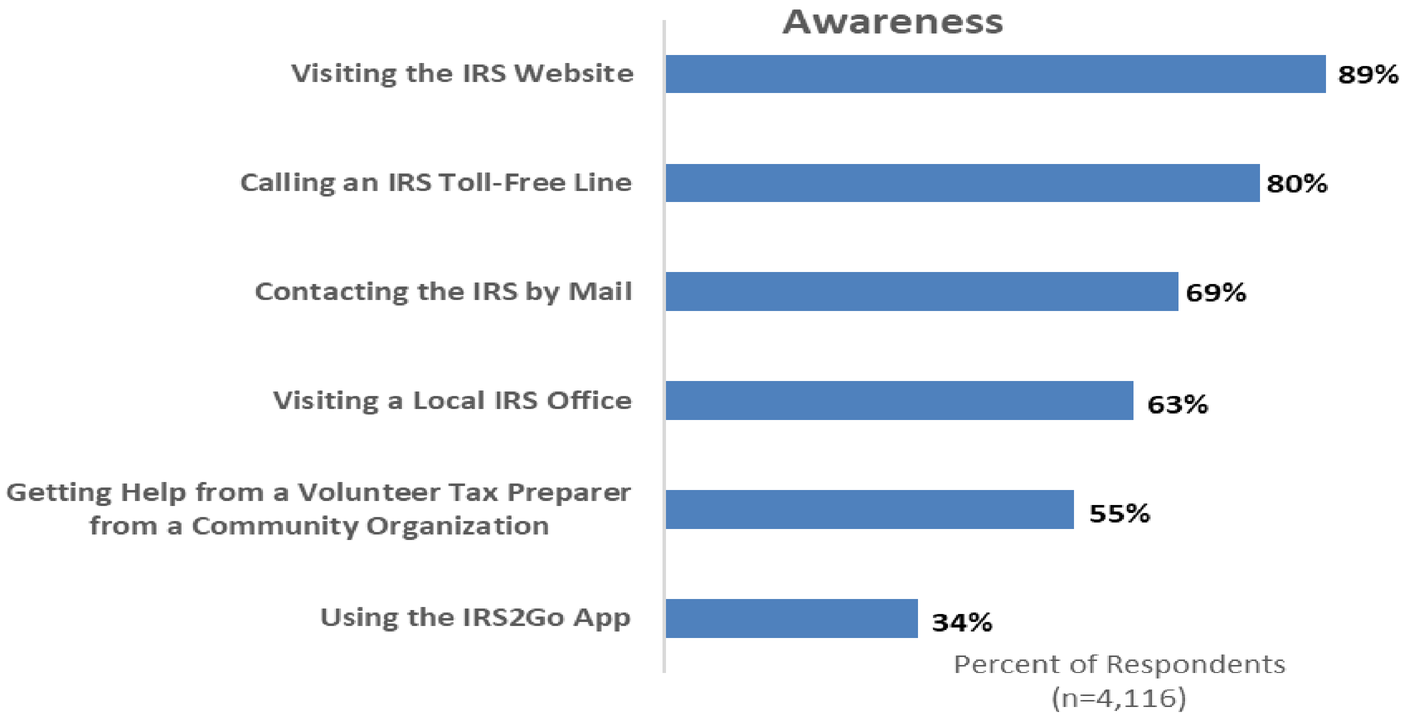
Major improvement over prior work

- Previous approach selected on taxpayers who contacted IRS → selection concerns
- Qualitative insights about what worked and what didn't
    - Focus groups → small N, but helpful to interpret statistics
- Insights for specific groups (rather than aggregating), like gig workers

Contrasting Messages

- Survey results → most obligations not too hard to understand
- Focus groups → lots of confusing elements, hard to get assistance with online tools, etc

Awareness rates were surprisingly high



Selection Concerns

- Hypothesis: survey-takers more likely to be interested in / enjoy tax preparation → more positive assessment
- Assessing selection
  - Compare demographics of respondents to true population
  - E.g., types of businesses, ages of filers, etc.
- More methodological details would be helpful
  - Survey response rate
  - Type of re-weighting? Based on which demographics
- Likely directions of bias

Future work?

- Compare perceptions to actual tax situation

- Randomized intervention to assess improvements to customer experience, like expedited phone number

Paper 3

Sun and Needham:
Federal Employment Tax Returns

Eligibility Versus Take-Up

- Straightforward to calculate take-up rate:

  Take-up = {Form 944 filers} / {Taxpayers eligible to file Form 944}

$$= \frac{43{,}786}{43{,}786 + 605{,}407} \approx 4.2\%$$

- Can compare take-up by industry and over time

- Could describe eligible but non-participating taxpayers

Hypotheses for Declining Take-up

- Eligibility tied to IRS pre-approval. Maybe response times went up / harder to get approval with funding cuts?

- Initial publicity from program’s introduction increased awareness

- Temporary regulations creating program suggested IRS would conduct outreach to each employer eligible to participate.

Policy Reforms?

- Raise eligibility threshold and index to inflation
- Remove requirement for prior approval
  - Delays undermine the value of the benefit for new companies
  - Pre-approval is burdensome for IRS
  - Could simply tie penalty to failure to file quarterly when tax liability > $1000
- Safe-harbor to qualify for annual filing based on prior year tax liability
- Seems like no new legislation would be needed … change the regs!

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

Session 4. Enhancing Taxpayer Customer Experience

Moderator:	***Christine Oehlert*** *IRS: RAAS*
Increasing Take-up of the American Opportunity Tax Credit	***Anne Herlache*** *IRS: RAAS*
Customer Experience for Small Business and Self-Employed Taxpayers	***Kahoa Bonhomme*** *IRS: SB/SE*
Are Annual Federal Employment Tax Returns Effective? An Economic Analysis of Filing, Reporting, and Payment Compliance Associated with Forms 943 and 944	***Yan Sun*** *IRS: RAAS*
Discussant:	***Jacob Goldin*** *Stanford Law School*

Research, Applied Analytics, and Statistics

TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

11th Annual IRS/TPC Joint Research Conference on Tax Administration

Wrap-Up

Eric Toder

Institute Fellow, Urban Institute, and
Codirector, Urban-Brookings Tax Policy Center