Changes to Voluntary Compliance Following Random Audits on Income Tax Returns

Murat Besnek & Allan Partington

Terminology

- **Compliant:** Taxpayers that didn’t require any adjustments to their net tax amounts during the random enquiry program (REP).
- **Non-Compliant:** Taxpayers that had some adjustments made to their net tax amounts during the REP.
- **Individuals Not in Business (INIB):** Taxpayers that are a part of the individuals tax gap population.
- **Small Business – Individuals in Business (SB-IIB):** Individual taxpayers that are a part of the small business tax gap population.
- **Small Business – Small Company (SB-SC):** Company taxpayers that are a part of the small business tax gap population.

Types of Revenue Collected from Audits

AUDIT YIELD	DIRECT FLOW-ON	INDIRECT FLOW-ON
This is the revenue collected from the adjustments made during the audit process.	This is referred to as the **direct deterrent effect** and includes the revenue collected from the changes in future voluntary compliance of audited taxpayers.	This is referred to as the **indirect deterrent effect** and includes the revenue collected from the spill-over effects on non-audited taxpayers.

Application to the ATO

- Activities like audits are the primary reason for compliance in the tax system.
- Without these strategies no rational taxpayer would comply, and instead prefer to free-ride.
- We know that it is not financially possible to pursue every taxpayer.
- So the payoffs for non-compliance is an expected value—the payoffs multiplied by the probability of not being caught.
- Taxpayer characteristics like being risk-averse could also impact these payoffs.
- The more credible the threat of an audit, the lower the payoffs for non-compliance, making it more beneficial to contribute rather than to free-ride.

REP Dataset

- The REP involves reviewing the returns of randomly selected taxpayers from the INIB, SB-IIB and SB-SC populations.
- At the commencement of this study, there were three years (2015, 2016 and 2017) of REP data available.
- The REP taxpayers of each year are analysed separately, but a joined estimate will also be provided.
- We only include REP taxpayers that have been contacted by the ATO, using the allocation date as a proxy for the date the taxpayer was contacted.

REP Dataset Cont.

- For each year of the REP, we also randomly select a control group that is approximately ten times larger in total numbers.
- We use net tax as our dependent variable (y_i).
- Once the dataset has been pulled together, we acquire all the net tax amounts for each taxpayer between the years of 2011-2020.
- We checked that the control group does not include any taxpayers that were contacted by the ATO during this period.
- A Wilcoxon rank-sum test is applied to confirm that the behaviours of the treated and control groups were similar in the pre-audit periods.

REP Dataset Cont.

- We remove all the net tax amounts that were part of the REP year, since we are only interested in voluntary compliance.
- For instance, for the taxpayers that were a part of the 2015 REP, we remove their 2015 returns from the dataset.
- This applies to both the treated and the control groups.

Audit Yield

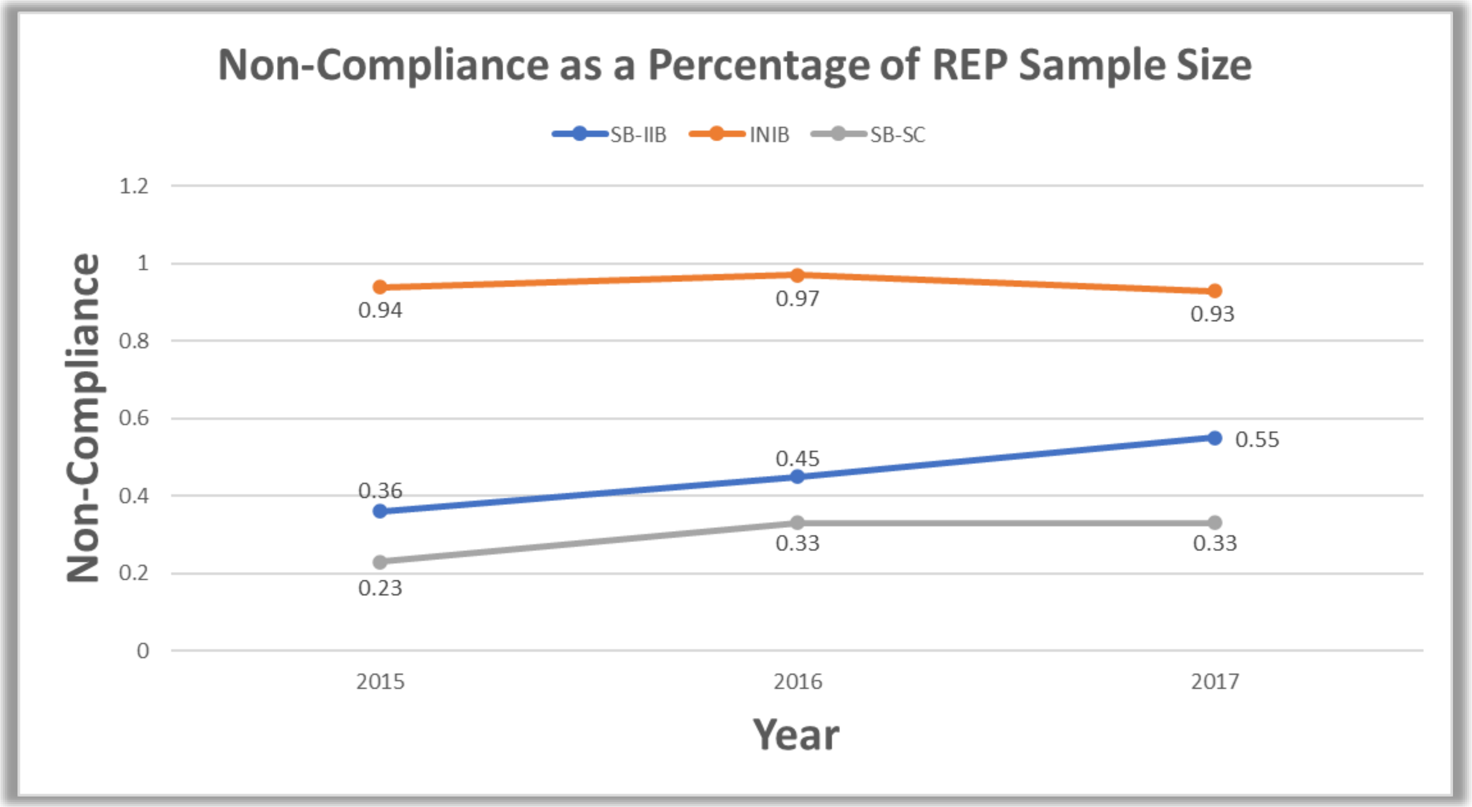
AVERAGE AUDIT YIELD

	2015	2016	2017	Joined
INIB	$1,071	$1,098	$881	$1,018
SB-IIB	$3,914	$2,001	$12,253	$6,936
SB-SC	$900	$2,705	$4,129	$2,433
REP Average	$1,962	$1,935	$5,754	$3,462

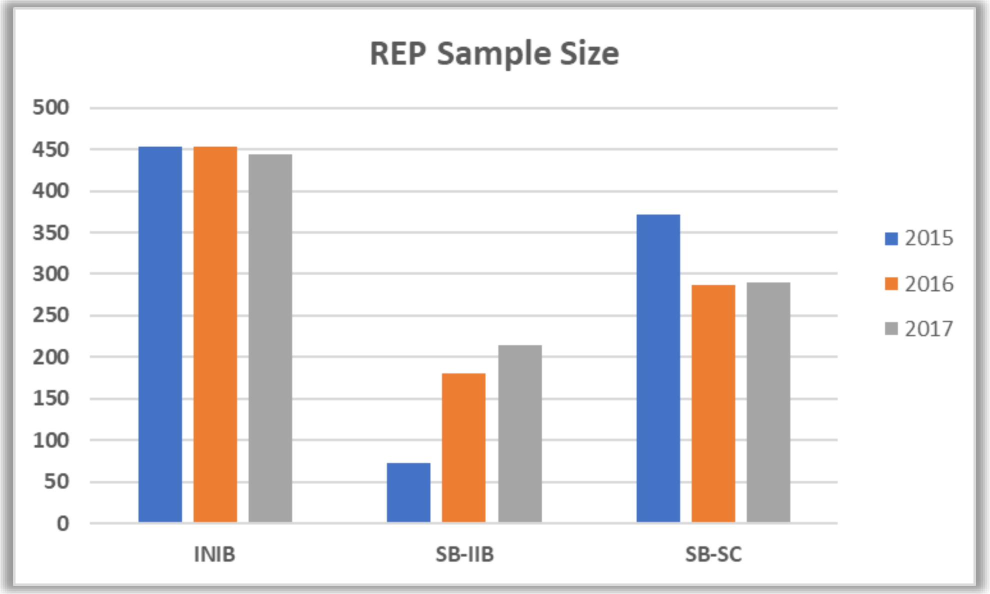
TOTAL AUDIT YIELD

	2015	2016	2017	Joined
INIB	$487,284	$497,386	$391,127	$1,375,797
SB-IIB	$301,397	$368,150	$2,658,827	$3,328,374
SB-SC	$334,933	$776,270	$1,197,467	$2,308,670
REP Total	$1,123,614	$1,641,806	$4,247,421	$7,012,841

Non-Compliance as a Percentage of REP Sample Size



REP Sample Size



Popular Fix: Using a Log-Linear Model

- Research in tax administration uses positively skewed datasets with $y_i = 0$ frequently.
- OLS estimators are not appropriate for statistical inference due to the violation of the normality assumption.
- The common approach is to estimate models using log-transformed dependent variables to deal with normality issues.
- The log-linear model is not suitable due to Jenny's inequality.
- Jenny's inequality implies that $E(\ln(y)) \neq \ln E(y)$, so retransforming coefficients from the log-linear model back to unlogged terms results in biased estimates (Motta, 2019).
- The retransformed estimates need to be adjusted for heteroscedasticity.
- If not, we can draw misleading conclusions about the parameters.

Adding a Positive Constant

- Another major issue with the popular fix is the inability to log zeros.
- So if we decide to use a log-log transformation to avoid un-logging the coefficient estimates, we still need to add a positive constant to all observations of y_i for the log-log transformation to be feasible.
- So deleting the zeros or giving them a small positive value can worsen the heteroscedasticity across the regressors (Motta, 2019).
- Moreover, the size of the positive constant needed will depend on the data at hand, so adding the smallest possible value (for example, the value of 1) is not the least harmful choice.
- In Bellégo et al. (2021), it is shown that the best value for the positive constant is not necessarily small nor equal to 1 contrary to common belief.

Estimating with PPML

- Instead of trying to correct for biasedness in log-linear or log-log models, the Poisson pseudo-maximum likelihood (PPML) is a robust substitute (Silva and Tenreyro, 2006; Correia, Guimarães and Zylkin, 2019).
- PPML is a method based on the Poisson regression with robust standard errors.
- The estimator is based on the conditional mean; therefore, the data does not have to have a Poisson distribution nor does y_i need to be an integer (Gourieroux, Monfort & Trognon, 1984).
- However with continuous data the assumption about the conditional mean equalling the conditional variance is unlikely to hold.
- For this reason the standard errors need to be based on the Eicker-Huber-White robust covariance estimator (Eicker, 1960; White, 1980).

Estimating with PPML Cont.

- The PPML model is becoming the industry standard in estimating multiplicative models for continuous data (following the advice of experts like Jeffery Wooldridge).
- The reason why the estimator is becoming popular is that the only condition required for consistency is the correct specification of the conditional mean.
- The estimator does not assume equality between the mean and the variance, nor does it require a constant variance.
- Poisson regression can also handle zeros in the dataset unlike the log-linear or log-log models that require the researcher to add a positive constant to all observations of y_i which may arbitrarily bias the estimates and their standard errors.

Difference in Differences Method

- Difference in differences (DID) method will be employed to measure the changes to voluntary compliance following the REP.
- D1 is the difference in net tax prior to the audit with those after the audit for the REP taxpayers .
- Any value in D1 can be a result of the REP, but also other possible events.
- To take into account some of these other possible events, we randomly select other taxpayers from the same population and year that were not a part of the REP.
- We also make sure that these taxpayers were not contacted by the ATO for other reasons during the period of interest, which is between 2011-2020.

DID Method Cont.

- D2 is the difference in net tax prior to the audit with those after the audit for these randomly selected taxpayers.
- Taking away D2 from D1 gives us the standard DID results.

DID = D1 – D2

$$(\mu_{t,post} - \mu_{t,pre}) - (\mu_{c,post} - \mu_{c,pre})$$

DID Method Cont.

- The standard DID method in regression form:

$$\text{NET TAX} = \beta_0 + \beta_1 DUMMY_{POST\_AUDIT} + \beta_2 DUMMY_{TREATED} + \beta_3 DUMMY_{POST\_AUDIT} DUMMY_{TREATED} + e_t$$

$$\beta_0 = \mu_{c,pre}$$

$$\beta_0 + \beta_1 = \mu_{c,post}$$

$$\beta_0 + \beta_2 = \mu_{t,pre}$$

$$\beta_0 + \beta_1 + \beta_2 + \beta_3 = \mu_{c,post}$$

DID Method Employed in Gemmell and Ratto (2012)

- Gemmell and Ratto (2012) introduce a new variation of the DID method to account for the differences between the behaviour of compliant and non-compliant taxpayers.

$$\text{NET TAX} = \beta_0 + \beta_1 DUMMY_{POST\_AUDIT} + \beta_2 DUMMY_{POST\_AUDIT} DUMMY_{COMPLIANT} + \beta_3 DUMMY_{POST\_AUDIT} DUMMY_{NON\_COMPLIANT} + \delta_t + e_t$$

$$\beta_0 + \beta_1 = \mu_{control,post}$$

$$\beta_0 + \beta_1 + \beta_2 = \mu_{compliant,post}$$

$$\beta_0 + \beta_1 + \beta_3 = \mu_{non\_compliant,post}$$

$$\delta_t = Individual\ fixed\ effects$$

DIRECT FLOW-ON EFFECTS FOR INIB USING THE PPML ESTIMATION METHOD

The average audit yields were $1,071 in 2015, $1,098 in 2016 and $881 in 2017 for this population. If we use the joined results, the average audit yields equal $1,018.

The per year direct flow-on effects for non-compliant taxpayers were $1,043 in 2015, -$2,740 in 2016 and $542 in 2017. If we use the joined regression coefficients, the per year direct flow-on effects for non-compliant taxpayers equal -$475.

		WRE PPML	COEFFICIENTS	P-VALUE
	POST-AUDIT	-	0.287	0.000
2015	POST-AUDIT * COMPLIANT	-	0.050	0.792
	POST-AUDIT * NON-COMPLIANT	$1,043	0.079	0.008
	POST-AUDIT	-	0.309	0.000
2016	POST-AUDIT * COMPLIANT	-	0.177	0.249
	POST-AUDIT * NON-COMPLIANT	-$2,740	-0.189	0.000
	POST-AUDIT	-	0.322	0.000
2017	POST-AUDIT * COMPLIANT	-	-0.160	0.199
	POST-AUDIT * NON-COMPLIANT	$543	0.045	0.054
	POST-AUDIT	-	0.307	0.000
JOINED	POST-AUDIT * COMPLIANT	-	-0.006	0.942
	POST-AUDIT * NON-COMPLIANT	-$475	-0.036	0.073

DIRECT FLOW-ON EFFECTS FOR SB-IIB USING THE PPML ESTIMATION METHOD

The average audit yields were $3,914 in 2015, $2,001 in 2016 and $12,253 in 2017 for this population. If we use the joined results, the average audit yields equal $6,936.

The per year direct flow-on effects for compliant taxpayers were -$2,720 in 2016. As for non-compliant taxpayers they were $3,077 in 2015 and $5,554 in 2016. If we use the joined regression coefficients, the per year direct flow-on effects for compliant taxpayers equal -$1,898 and for non-compliant $2,616.

		WRE PPML	COEFFICIENTS	P-VALUE
2015	POST-AUDIT	-	0.264	0.000
	POST-AUDIT * COMPLIANT	-	-0.118	0.408
	POST-AUDIT * NON-COMPLIANT	$3,077	0.292	0.016
2016	POST-AUDIT	-	0.217	0.000
	POST-AUDIT * COMPLIANT	-$2,720	-0.193	0.019
	POST-AUDIT * NON-COMPLIANT	$5,554	0.394	0.000
2017	POST-AUDIT	-	0.269	0.000
	POST-AUDIT * COMPLIANT	-	-0.094	0.224
	POST-AUDIT * NON-COMPLIANT	-	0.032	0.723
JOINED	POST-AUDIT	-	0.248	0.000
	POST-AUDIT * COMPLIANT	-$1,898	-0.148	0.007
	POST-AUDIT * NON-COMPLIANT	$2,616	0.204	0.002

DIRECT FLOW-ON EFFECTS FOR SB-SC USING THE PPML ESTIMATION METHOD

The average audit yields were $900 in 2015, $2,705 in 2016 and $4,129 in 2017 for this population. If we use the joined results, the average audit yields equal $2,433.

The per year direct flow-on effects for compliant taxpayers were $3,742 in 2015, $4,981 in 2016 and $5,529 in 2017. As for non-compliant taxpayers they were $18,130 in 2016. If we use the joined regression coefficients, the per year direct flow-on effects for compliant taxpayers equal $4,848 and for non-compliant $5,955.

		WRE PPML	COEFFICIENTS	P-VALUE
2015	POST-AUDIT	-	0.164	0.000
	POST-AUDIT * COMPLIANT	$3,742	0.189	0.100
	POST-AUDIT * NON-COMPLIANT	-	0.001	0.995
2016	POST-AUDIT	-	0.120	0.000
	POST-AUDIT * COMPLIANT	$4,981	0.200	0.023
	POST-AUDIT * NON-COMPLIANT	$18,130	0.728	0.004
2017	POST-AUDIT	-	0.185	0
	POST-AUDIT * COMPLIANT	$5,529	0.195	0.008
	POST-AUDIT * NON-COMPLIANT	-	0.056	0.651
JOINED	POST-AUDIT	-	0.154	0.000
	POST-AUDIT * COMPLIANT	$4,848	0.197	0.000
	POST-AUDIT * NON-COMPLIANT	$5,955	0.242	0.039

Conclusions

- We used a random dataset to improve the accuracy of the estimates.
- We employed the industry standard when it came to the modelling phase, that being the PPML method (following the advice of Jeffery Wooldridge and many other academic papers).
- Our approach/model does not deviate from what the raw data suggests (other than making the estimates more precise), which can be confirmed by comparing it to the standard DID estimates which only require algebra to compute.
- The direct flow-on effect for non-compliant taxpayers in the INIB population is negative.
- The direct flow-on effect for compliant taxpayers in the SB-IIB population is negative.
- The direct flow-on effect for non-compliant taxpayers in the SB-IIB population is positive.

Conclusions Cont.

- The direct flow-on effect for both compliant and non-compliant taxpayers in the SB-SC population is positive, but larger for non-compliant taxpayers.
- Yearly treatment effects seem to remain steady, lasting multiple years following the audit allocation date.

What’s Next?

- Extend the analysis using 2018 REP data.
- Incorporate the operational audit data to see if it suggests that risk based audited taxpayers behave differently to the taxpayers in the REP.
- Attempt to estimate the indirect flow-on effect using ATO data.

The Long-Term Impacts of Audits on Nonfiling Taxpayers

IRS-TPC Research Conference on Tax Administration

June 22, 2023

India Lindsay and Jess Grana (The MITRE Corporation)
Alan Plumley (IRS)

Research, Applied Analytics & Statistics
KNOWLEDGE DEVELOPMENT & APPLICATION

SOLVING PROBLEMS FOR A SAFER WORLD

Background: Nonfiling Taxpayers

- Nonfilers are responsible for $32 billion (9%) of the individual income tax gap[1]
  - Population of interest: nonfilers with at least $100k income
    - Income determined from 3rd party reported income
  - Higher earning nonfilers owe greater than 73% of the nonfiling gap[2]

TY 2014-2016 Estimates of Tax Gap

		Gross Tax Gap					
Total True Tax Liability	Tax Paid Voluntarily & Timely	Nonfiling	Underreporting	Under-payment	Gross Tax Gap	Enforced & Other Late Payments	Net Tax Gap (Tax Not Collected)
$3,307	$2,811	$39	+ $398	+ $59	= $496	- $68	= $428
By Type of Tax							
Individual Income Tax	Individual Income Tax	Individual Income Tax	Individual Income Tax	Individual Income Tax	Individual Income Tax	Individual Income Tax	Individual Income Tax
$1,740	$1,383	$32	+ $278	+ $47	= $357	- $51	= $306

Source: https://www.irs.gov/pub/irs-pdf/p1415.pdf

1. https://www.irs.gov/pub/irs-pdf/p1415.pdf
2. TIGTA 2020

Background: Decline in Audit Resources

- Increase in the number of nonfilers identified every year
  - 7.5 million in 2010 → 10.7 million in 2016
- Decrease in resources to audit these individuals
  - 3.5 million cases started in 2010 → 0.8 million cases started in 2018
- IRS 2020 Nonfiler Enforcement Initiative[1] promises stronger pursuit of nonfilers, specifically higher earning individuals
- IRS Inflation Reduction Act Strategic Operating Plan[2] to “address high-dollar compliance issues”
- ROI metrics needed to evaluate indirect impact of audits

<u>Research Question:</u>

What is the effect of an audit on the long-term filing behavior of a nonfiling taxpayer?

1. https://www.irs.gov/newsroom/irs-increasing-focus-on-taxpayers-who-have-not-filed-tax-return
2. IRS 2023

Relevant Literature

- Conflicting theory on how enforcement affects future filing behavior:
  - Deterrence effect: audits deter future noncompliance
  - Bomb crater effect: individuals are more likely to be noncompliant following an audit
- Literature on nonfilers suggest factors influencing filing behavior include income visibility, persistence of filing behavior, and taxpayer's perception of government and sense of moral duty
  - Erard et al. (2022) is one of the first papers to consider higher earning nonfilers
- Limited studies on indirect effects of enforcement on nonfilers
  - Taglakis (2014) studied effect of audits in Greece; a 1% increase in number of audits leads to a 0.4% increase in direct revenue and 0.1% increase in indirect revenue for high wealth individuals and nonfilers
  - Datta et al. (2015) found Automated Substitute for Return activities increased likelihood of filing by 11, 21, and 27 percentage points in 2-4 years post treatment
- **Gap in literature analyzing both the behavior of higher earning nonfilers and role of IRS enforcement on future filing behavior**

Nonfiler Field Audit Selection Process

Case Creation Nonfiler Identification Process (CCNIP)
Individuals that have not filed but have reported income or a prior year return are identified by CCNIP

SBSE Field Audit Assignment
IRS estimates tax liability, assigns priority score, and assigns eligible nonfilers to Field personnel based upon available resources

01

02

03

04

Third Party Reporting
Employers, entities, and financial institutions report taxpayer compensation to IRS; compiled with prior year return information

Nonfiler Audit Selection
Selection filters are applied and nonfilers may be distributed to one of three enforcement functions: ASFR, Collection, or Field Exam

Selection process details primary procedure for identifying nonfilers for in-person field audits; nonfilers are identified for field audits via alternate processes

Sample Design

- Taxpayer data obtained from the IRS Compliance Data Warehouse (CDW)
- *Baseline year* = Tax Year (TY) the taxpayer entered the sample, due to audit or eligibility, between TY 2009-2014

Treatment Group

- Nonfilers **audited** under Field exam identified from examination records
- Excludes pickups

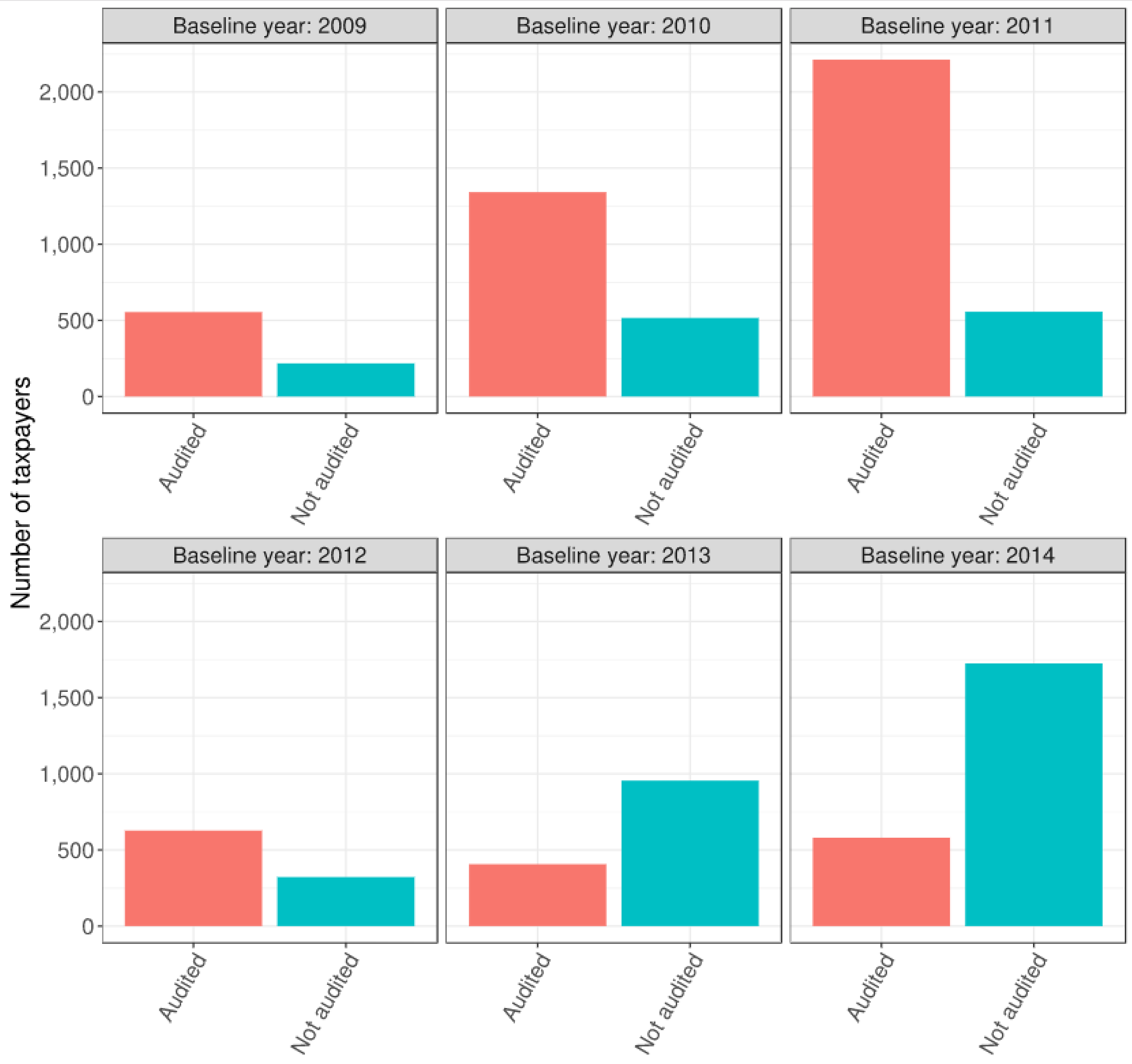
Control Group

- Nonfilers **eligible but unaudited** for Field exam
- Identified by replicating audit selection process
- Excludes late filers, secondary filers, and individuals filing in response to notices

Sample Design

Sample Size by Baseline Year

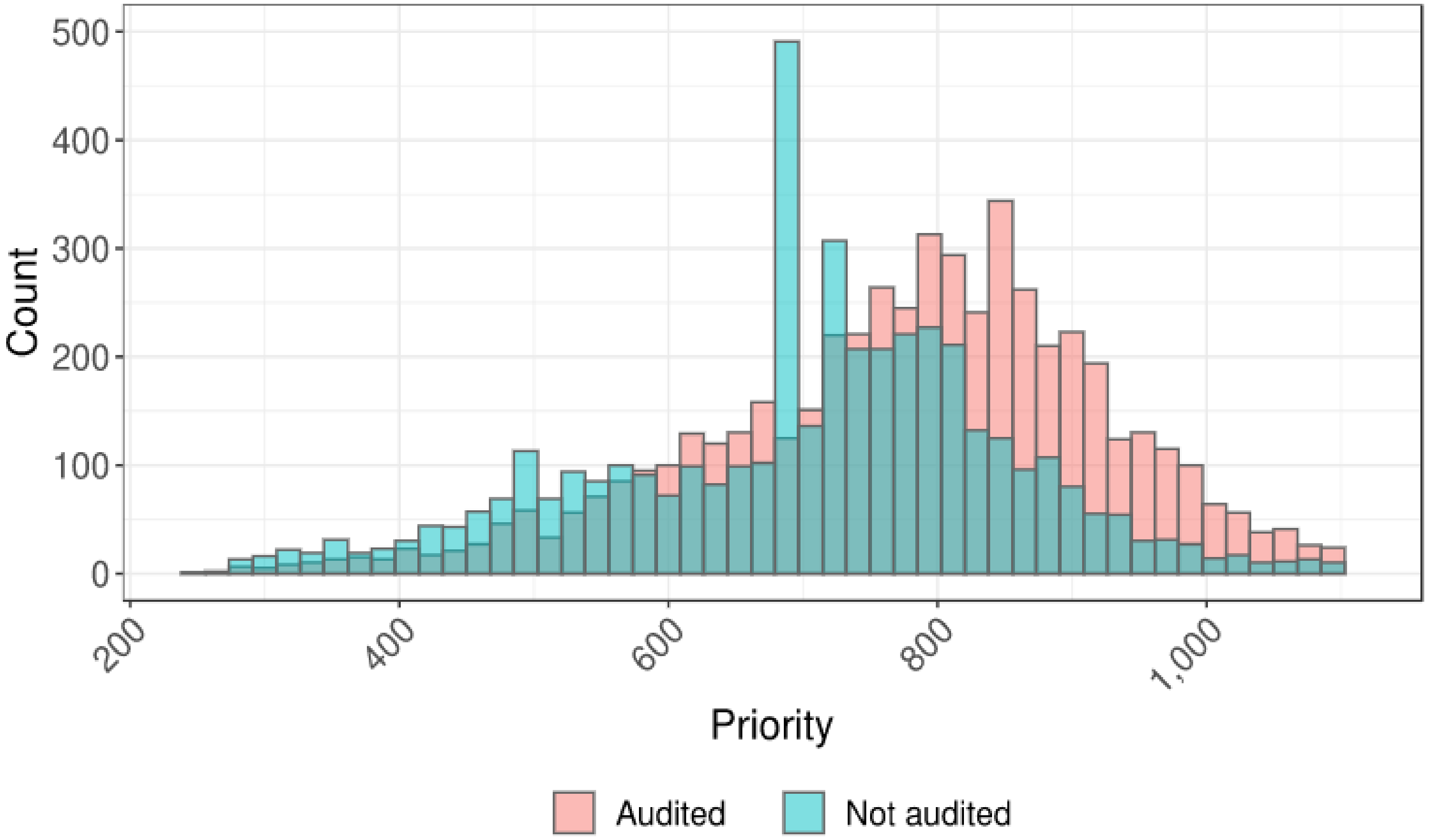
- Treatment group: 5,727
- Control Group: 4,297
- Fewer audits conducted after 2011 potentially due to changes in audit resources
- Dropped from sample if:
    - Deceased
    - Identified for audit via alternate procedure
    - In treatment group and missing examination record data
    - In control group and audited in 6 years surrounding baseline



Source: MITRE analysis of CDW data

Overlap in distribution of priority, across groups

- Priority is an IRS-internal metric ranking taxpayers for audit selection based upon balance due and likelihood of securing balance due

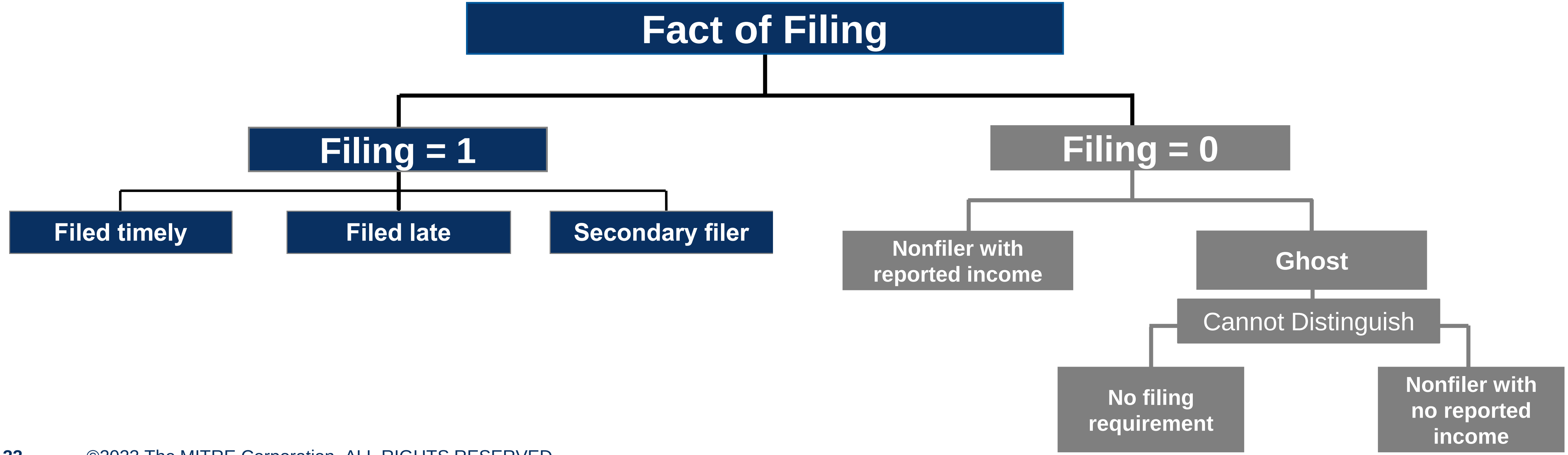


Source: MITRE analysis of CDW data

Dependent Variable

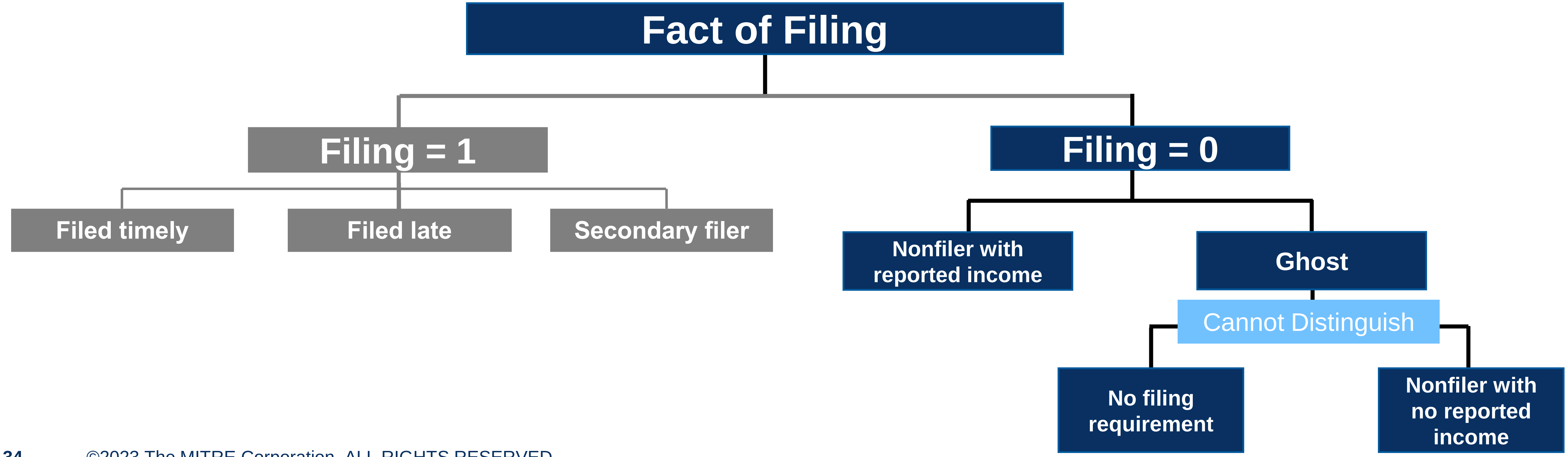
- Filing behavior for 5 TYs prior to and 8 TYs post baseline year

- Fact of Filing $= \begin{cases} 1, & taxpayer\ filed\ a\ return \\ 0, & otherwise \end{cases}$



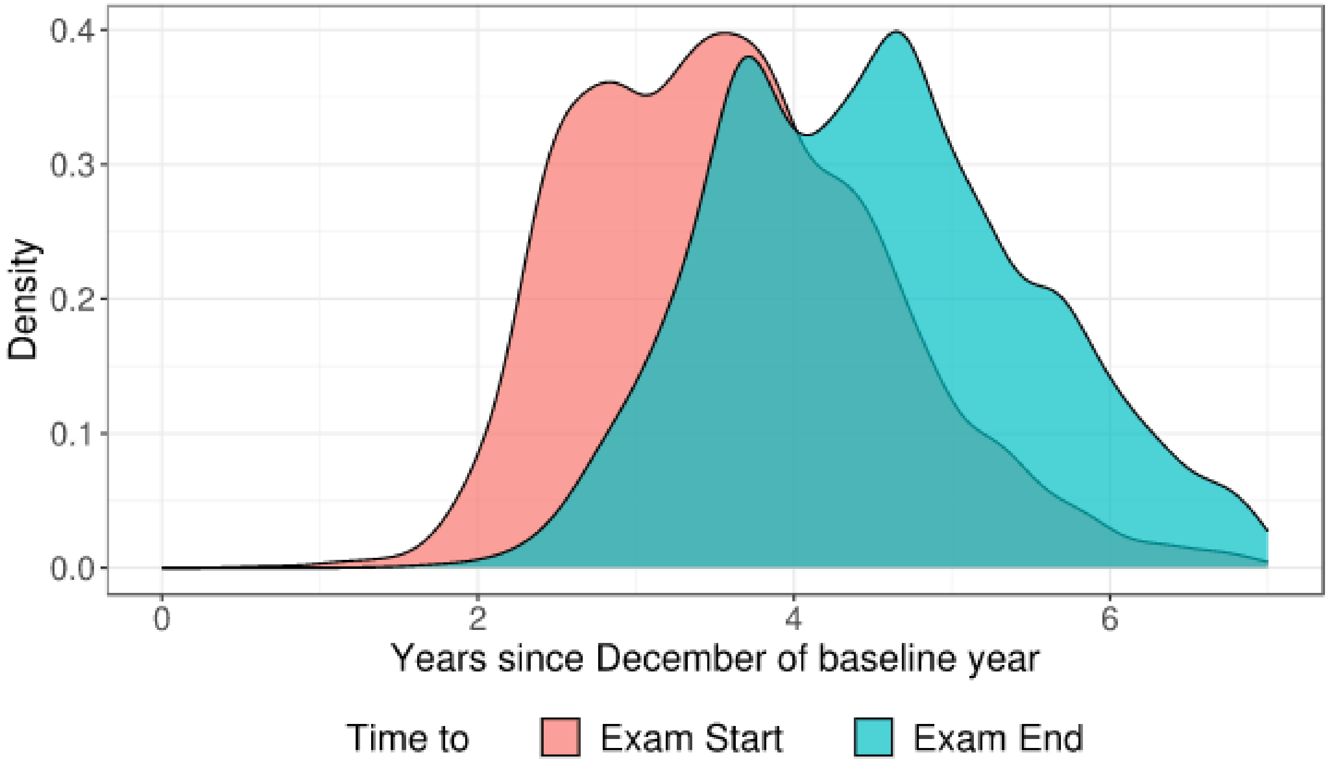
Dependent Variable

- Filing behavior for 5 TYs prior to and 8 TYs post baseline year
- Fact of Filing $= \begin{cases} 1, & taxpayer\ filed\ a\ return \\ 0, & otherwise \end{cases}$



Audit Timing

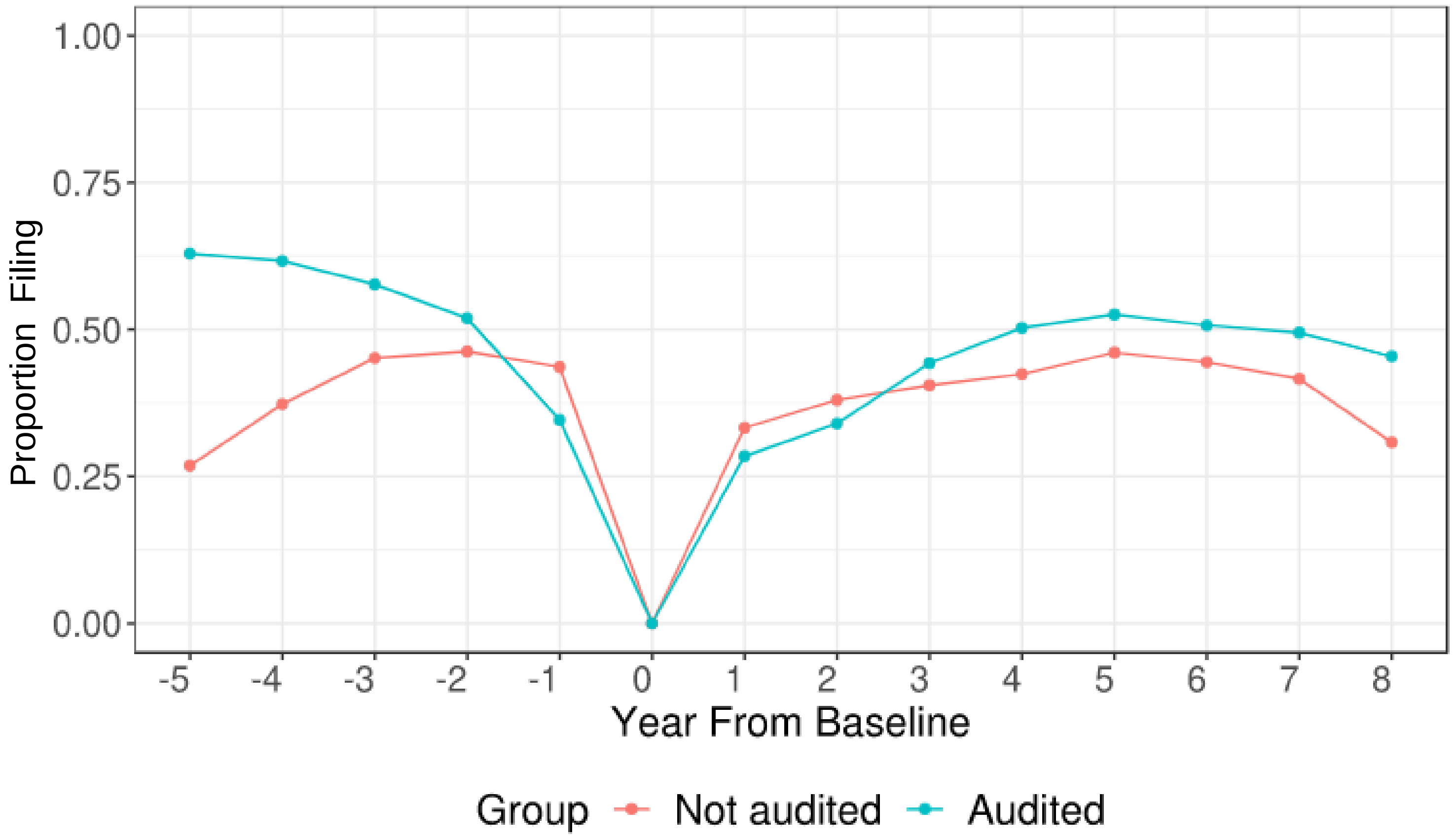
- Audits begin 2-5 years after baseline year
- Audits end 3-6 years after baseline year
- We hypothesize an indirect effect will not be observed until at least two years after baseline year



Source: MITRE analysis of CDW data

Filing over Time: Control vs Treatment

- All individuals are nonfilers in baseline year
- Baseline year interrupts patterns of filing behavior
- Audited taxpayers more likely to file post audit



Source: MITRE analysis of CDW data

Linear Probability Model

$$Fact\ of\ Filing_{it} = \beta_0 + \beta_1 Audit_i + \beta_{2-14} Year\ from\ Baseline_{it} + \beta_{15-27} Audit_i * Year\ from\ Baseline_{it} + \boldsymbol{\alpha Taxpayer\ Controls_i} + \boldsymbol{\tau Tax\ Year_t} + \epsilon_{ij}$$

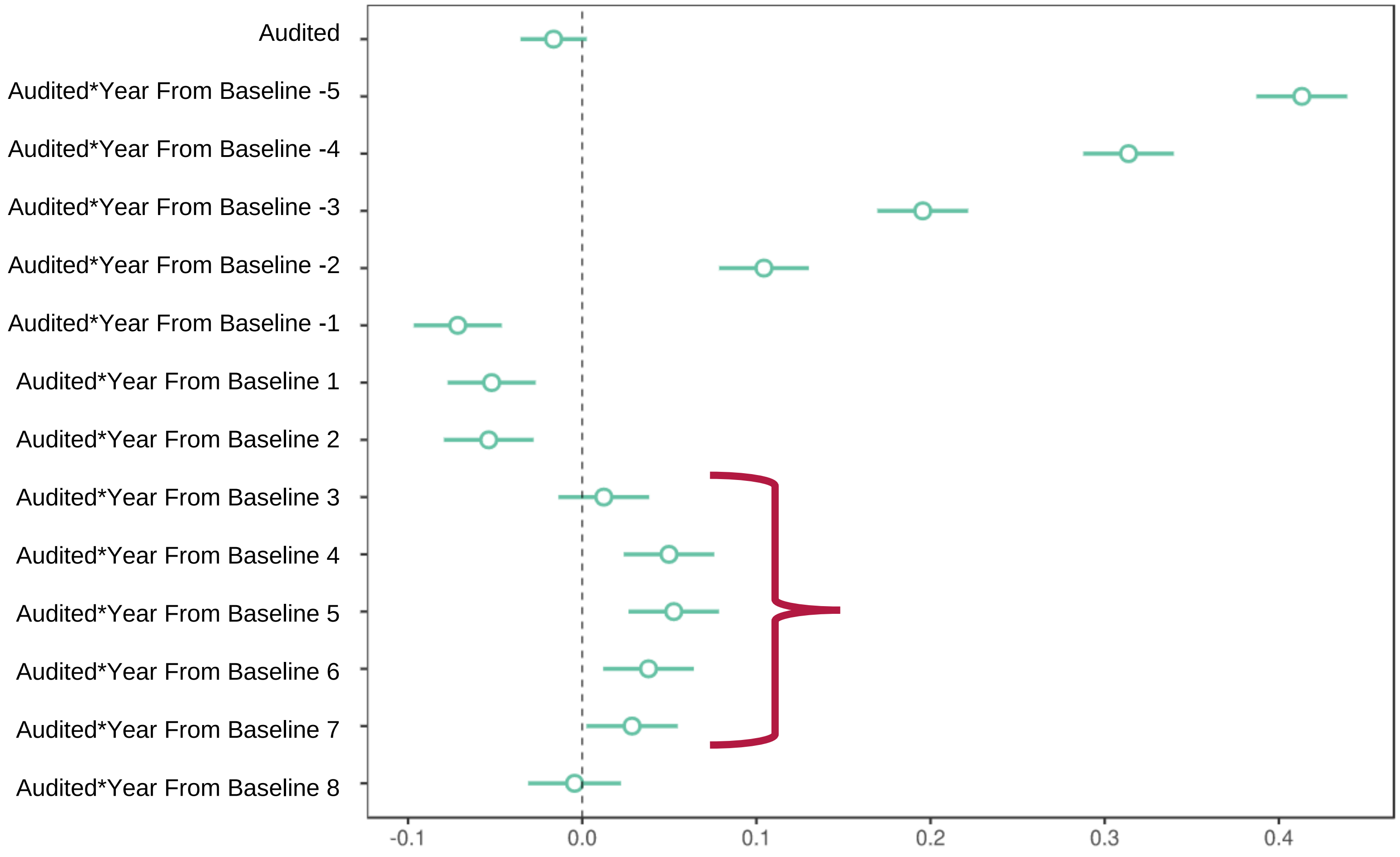
Where:

- ***Audit*** captures difference in average filing behavior across groups for all years
- ***Year from Baseline*** captures filing behavior for each of the 13 years surrounding baseline
- ***Audit * Year from Baseline*** captures indirect effects of an audit on filing behavior
- ***Taxpayer Controls*** are time-invariant, capturing demographic characteristics, financial characteristics, and past filing behavior in baseline year
- ***Tax Year*** is a set of fixed effects capturing yearly fluctuations across all taxpayers

Model Results: Indirect Effects

Indirect effect on filing behavior observed 4-7 years post baseline

- Audited group is 2.9-5.3% more likely to file in the 4-7 years from treatment
- Negative effect in years from baseline -1 through 2 suggest persistence of filing behavior for audited individuals surrounding year of noncompliance



Source: MITRE analysis of CDW data

Model Results: Control Variables

- Presence of visible income sources increases likelihood of filing (investment income has strongest effect at 9.6%)
- Residing in a state taxing individual income increases likelihood of filing by 17.9%
- Persistence of filing behavior
  - Taxpayers filing a return in prior year are 20.3% more likely to file
  - Taxpayers not present in IRS records in prior year are 10.7% less likely to file

Discussion

- Results support value of audits as a tool to encourage future filing in nonfilers
  - Audited taxpayers are 2.9-5.3% more likely to file in 4-7 years post treatment
  - Impact of an audit on future filing peaks 5 years after an audit, fades 7 years after

- Compared to estimated indirect effect of an ASFR on future compliance (Datta et al., 2015), indirect effect of a Field audit is smaller
  - ASFR increased likelihood of filing by 11%, 21%, and 27% in 2-4 years post treatment
  - Difference in estimates may be indicative of higher compliance rates in lower income populations

Limitations & Future Research

Limitations	Future Research
Indirect effects in terms of **revenue**?	Estimation of **total tax model** to obtain dollar-valued estimates
72% of audited group experienced **multiple** audits	Analysis of indirect effect on **single vs. multiple-audited taxpayers**
Assumption that **ghost taxpayers** have a filing obligation	**Sensitivity analysis** on ghost assumption; verification of tax liability in off-baseline years
Analysis **constrained by third-party reported data** only available for baseline year	Richer set of **time-varying control variables** (received notices, type of nonfilers); identifying data for ghosts

Thank you

Field Audit Selection Process

06

Audit Inventory
TDI inventory of nonfilers consists of individuals identified by CCNIP plus various referral programs

07

Audit Selection
Selection filters are applied and nonfilers may be distributed to one of three enforcement functions: ASFR, Collections, or **Field Exam**

08

Audit Allocation
Individuals identified for Field are sorted by region and Field offices, based upon available resources

09

Audit Prioritization
Individuals are prioritized based upon various factors, such as a function of the likelihood of securing a return and amount of balance due

10

Audit Assignment
Individual selected for field audit are assigned to field personnel based upon priority

Sample Cleaning

Dropped taxpayers that died within 8 years of baseline year (1,519)

Dropped taxpayers selected for audit by alternate procedures and not in CCNIP (1,811)

Dropped audited taxpayers in treatment group if:

Missing or unmatched examination data (677)	Filed F1040 late but prior to exam start (161)

Dropped eligible-but-unaudited taxpayers in control group if:

Audited 6 years prior to baseline (114)	Audited 6 years post baseline (95)

Deduplication rules (690):

If audited multiple times, first audit year assigned as baseline	If eligible-but-unaudited multiple times, first eligible year assigned as baseline	If audited and eligible, first audit year assigned as baseline

Independent Variables

Demographic Control Variables

- PY Filing status collapsed into two categories
- Majority of taxpayers between 30-65, have a single/other filing status, and reside in a state taxing individual income

Variable	Treatment Group	Control Group
Census Region		
East North Central	11%	8%
East South Central	7%	4%
Mid Atlantic	13%	13%
Mountain	7%	7%
New England	5%	4%
Pacific	15%	15%
South Atlantic	17%	22%
West North Central	5%	3%
West South Central	20%	15%
Not Available	1%	9%
Income Tax State	74%	74%
Over 65	4%	7%
Under 30	7%	12%
PY Filing Status		
Single/other	71%	88%
Married filing jointly	29%	12%
PY EITC	9%	3%

PY: Prior Year

Independent Variables

Financial Control Variables

- Total IRP income: sum of all reported income
- $100k threshold not enforced for treatment group (see appendix)
- Income difference: difference in income reported for current year from prior year
- Majority of treatment group have SE income
- Majority of control group have investment and/or other income

Variable	Treatment Group	Control Group
Total IRP Income	$551,114	$581,269
$100k Threshold Indicator	54%	100%
Number of IRP Forms	35	43
Income Difference from PY	$478,408	$533,118
SE Income	69%	46%
Investment Income	44%	69%
Retirement Income	20%	21%
Broker Transaction Income	19%	32%
Other Income	29%	59%

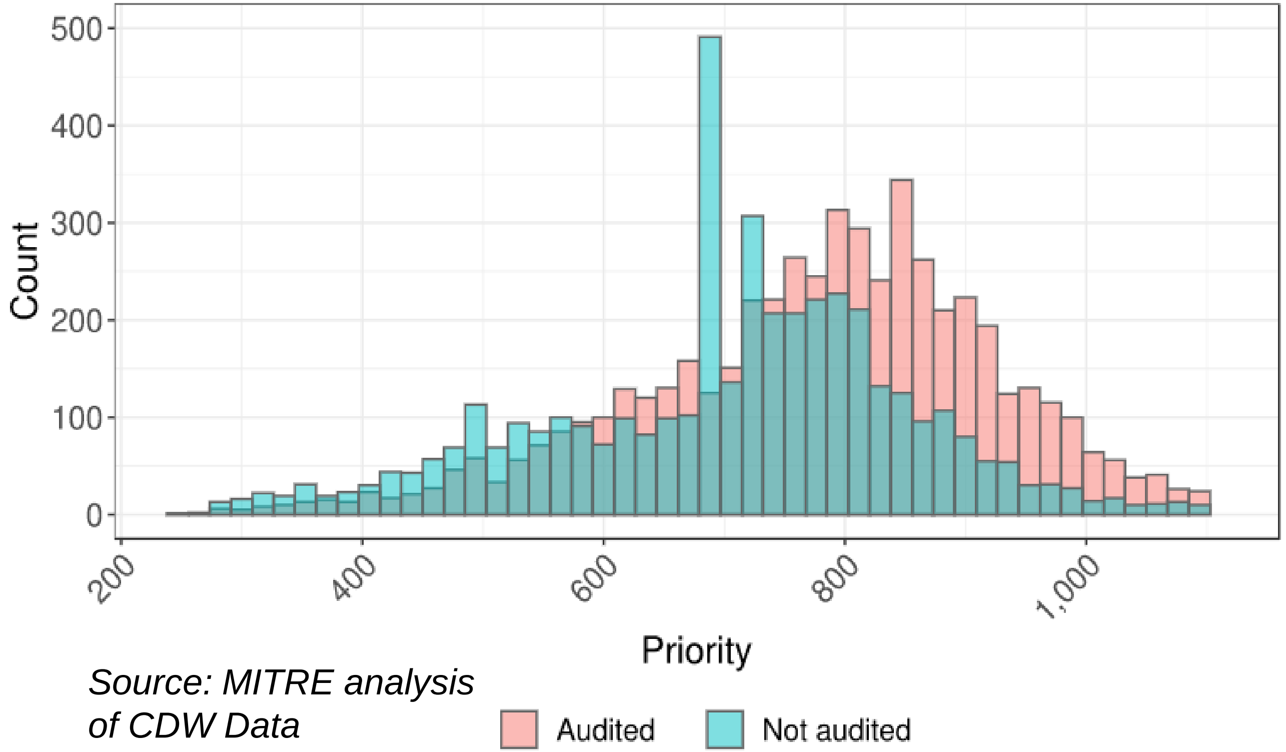
Dollar-denominated variables (Total IRP Income and Income Difference from PY) are expressed in terms of 2018 dollars. Other than Baseline Priority and Number of IRP forms, all other variables reflect percentages.

Independent Variables

Prior Filing Behavior Control Variables

- Priority is an IRS-internal metric ranking taxpayers for audit selection based upon balance due and likelihood of securing balance due
- Common support in priority scores across groups
- Majority of audited taxpayers experienced some type of audit in last 6 years

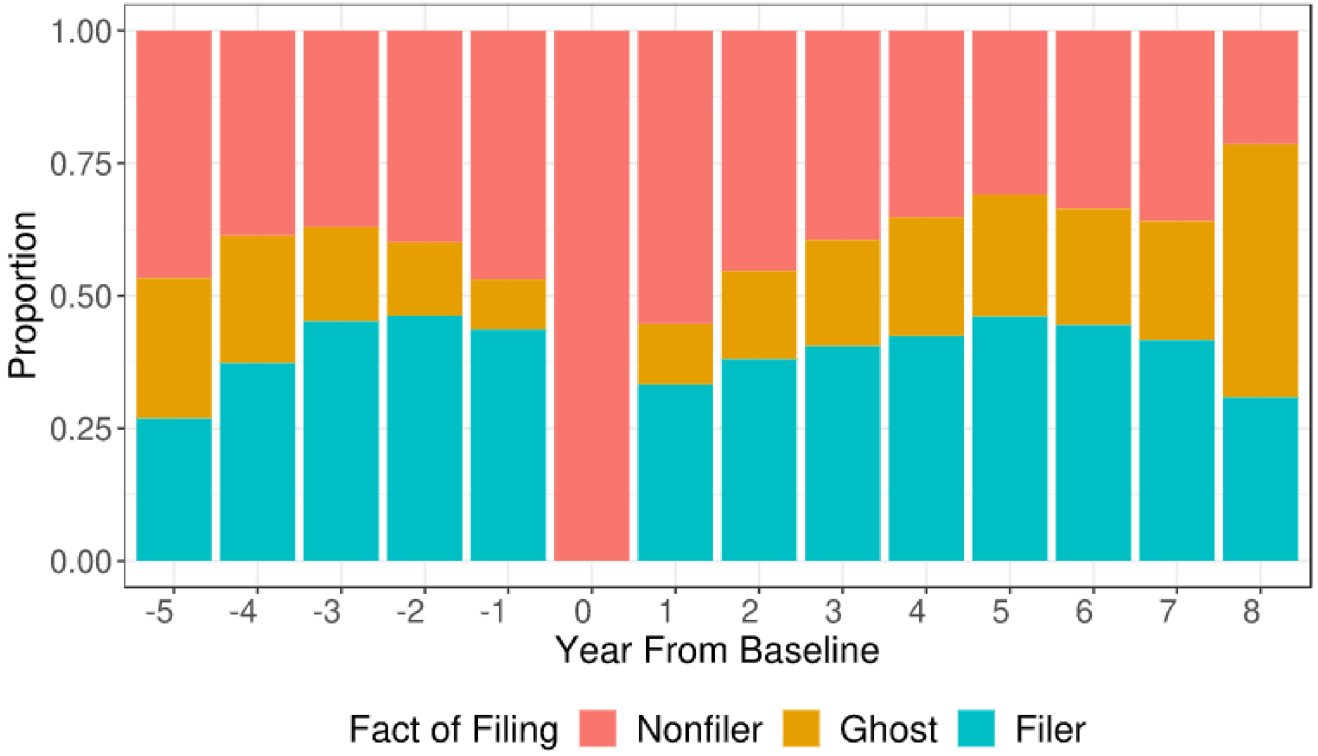
Variable	Treatment Group	Control Group
Filed in PY	78%	47%
Ghost in PY	0%	9%
Any Audit Last 6 TYs	53%	3%
Baseline Priority	813	712



Source: MITRE analysis of CDW Data

Distribution of Ghosts in Control Group

- Proportion of ghosts ranges from 9.5 to 47.7%
- All taxpayers in treatment group present in IRS records for years of interest

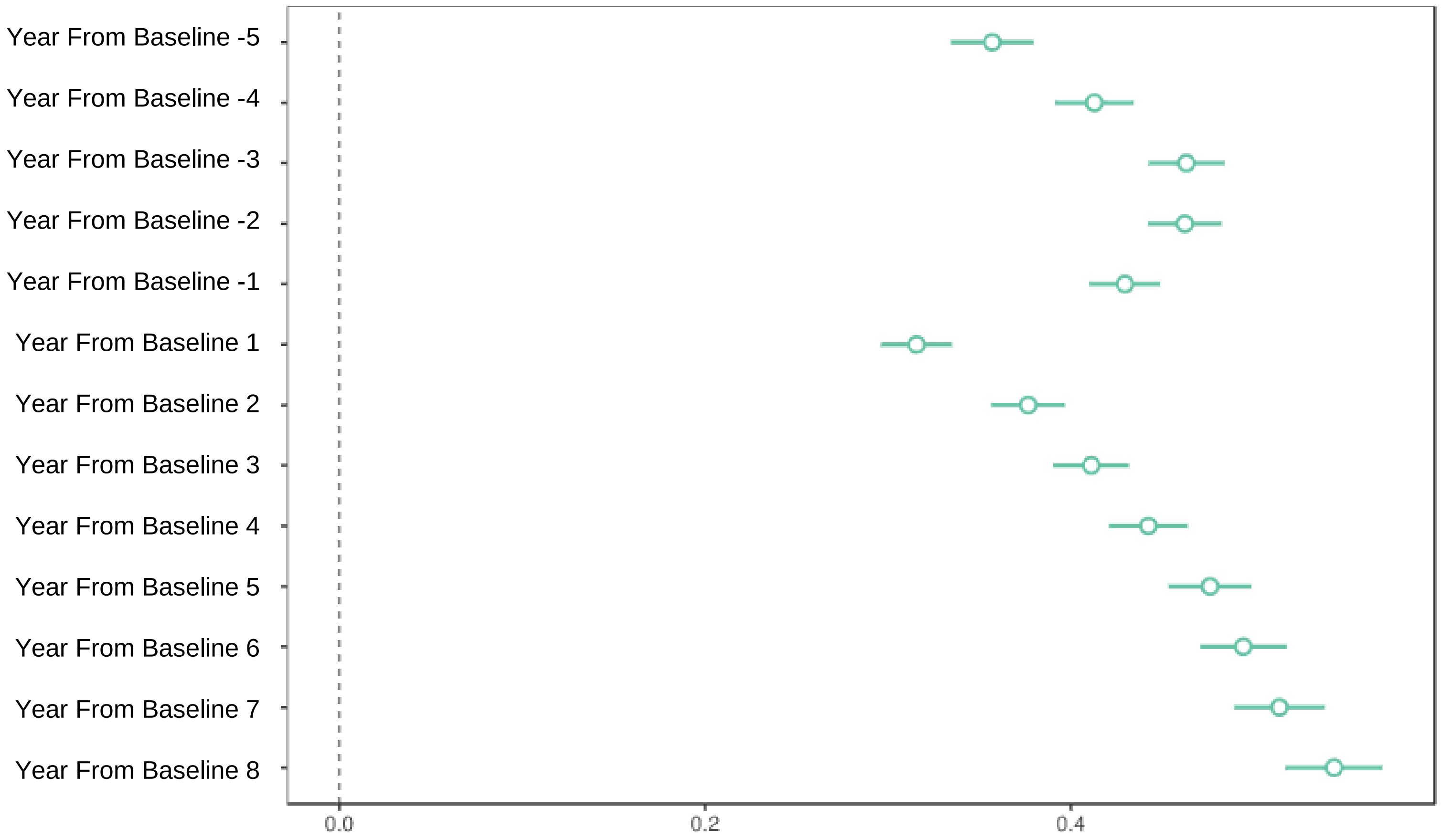


Source: MITRE analysis of CDW data

Model Results: Year From Baseline

All taxpayers more likely to file a return in off-baseline years

- In general, baseline year is an outlier year
- Pattern of decreased filing behavior leading up to baseline
- Patter of increased filing behavior after baseline

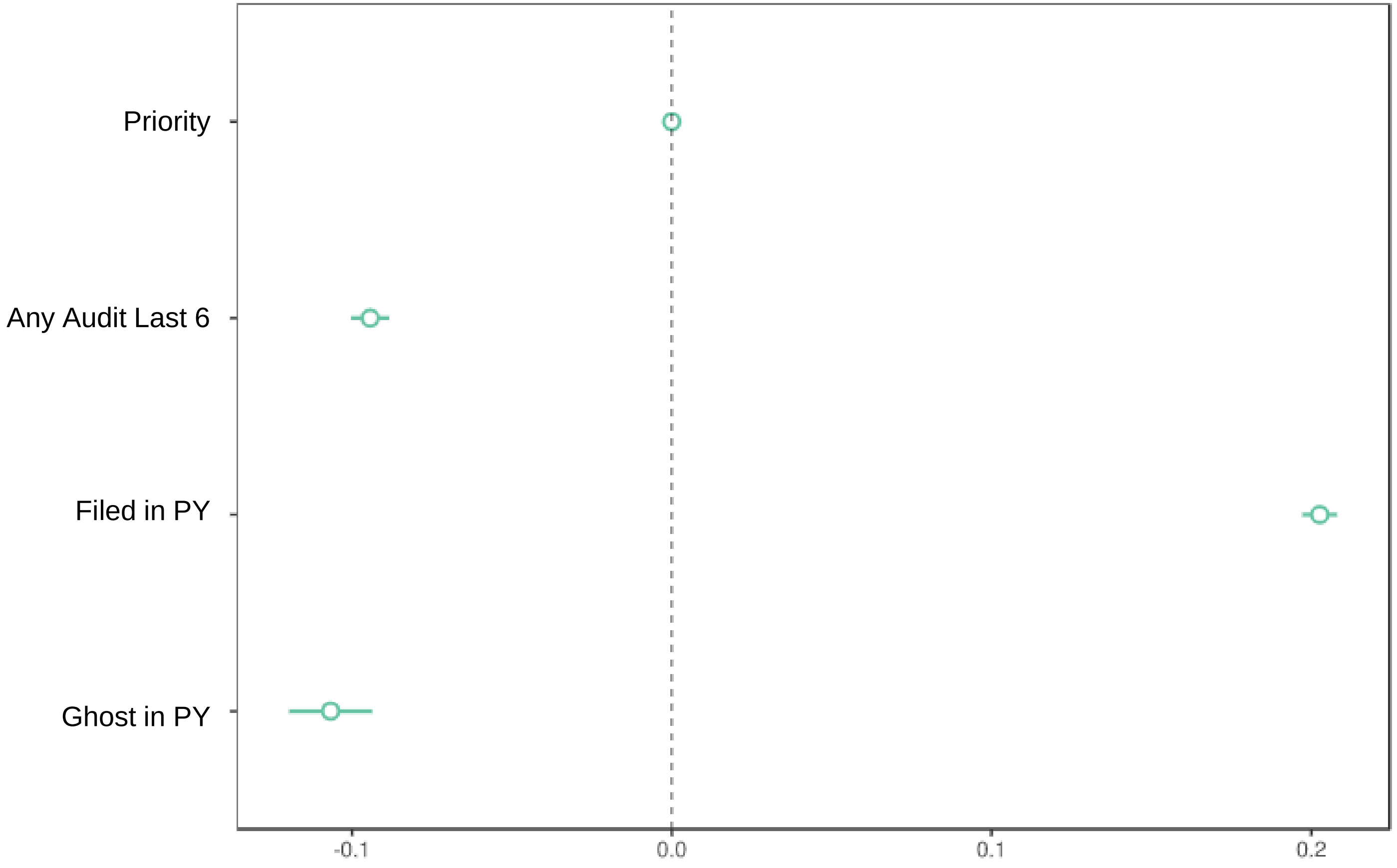


Source: MITRE analysis of CDW data

Model Results: Past Filing Behavior

Observed persistence in an individual's filing behavior

- Taxpayers filing a return in prior year are 20.3% more likely to file
- Taxpayers not present in IRS records in prior year are 10.7% less likely to file
- Any audit in 6 years prior to baseline reduces likelihood of filing by 9.4%
- Priority does not have a significant effect on probability of filing

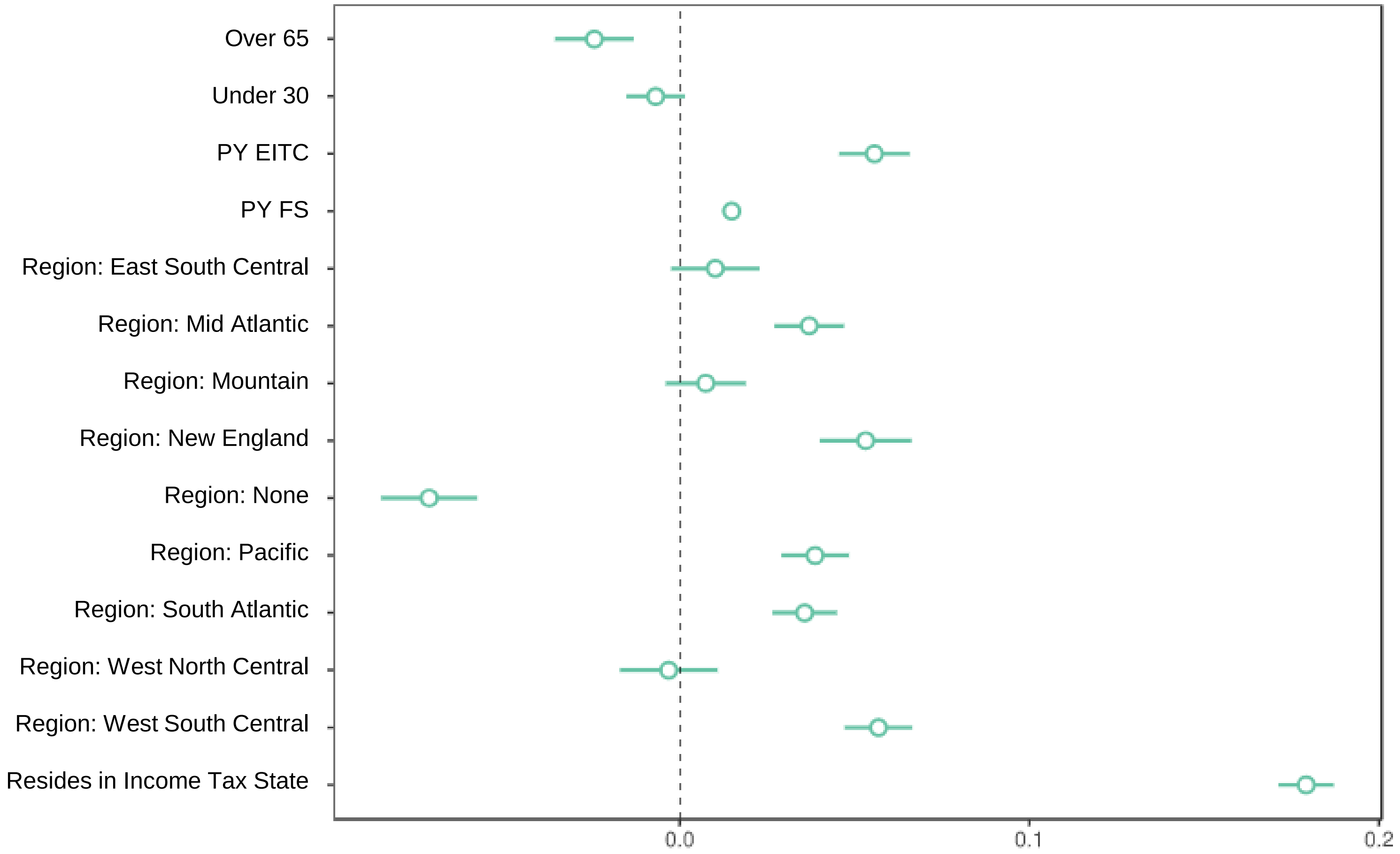


Source: MITRE analysis of CDW data

Model Results: Demographic Characteristics

Residence in a state taxing income has strongest influence

- Residing in a state taxing individual income increases likelihood of filing by 17.9%
- Otherwise filing behavior varies with geography
- Taxpayers over 65 are 2.4% less likely to file
- Taxpayers married filing jointly are 1.5% more likely to file

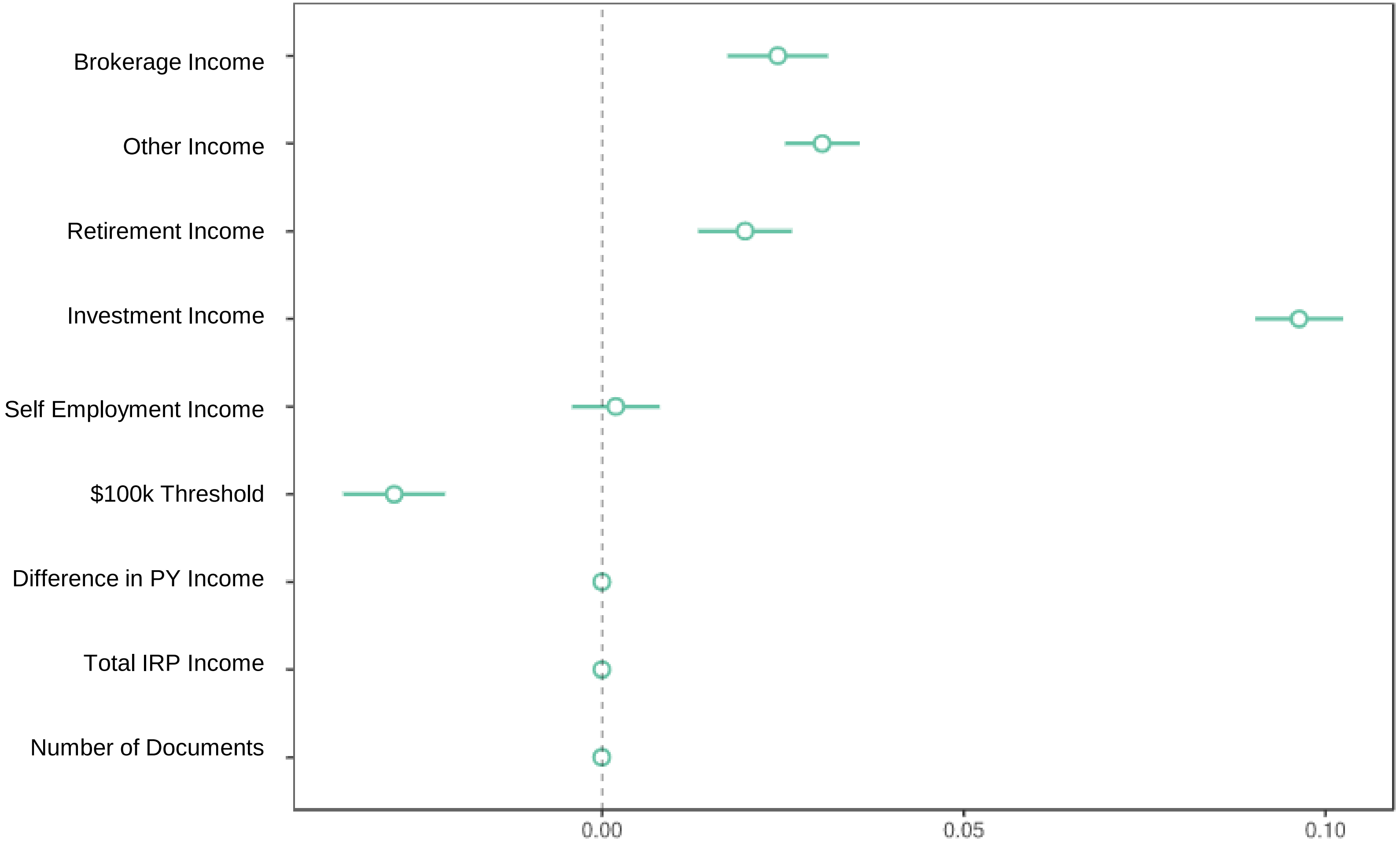


Source: MITRE analysis of CDW data. Region of comparison is East North Central

Model Results: Financial Characteristics

Presence of visible income sources increases likelihood of filing

- Investment income has strongest effect (9.6%)
- Significance of SE income may be obscured by measurement error
- Individuals earning greater than $100k are 2.9% less likely to file
- Actual amount of income insignificant
- For each additional document reported to the IRS, a taxpayer is 0.002% less likely to file

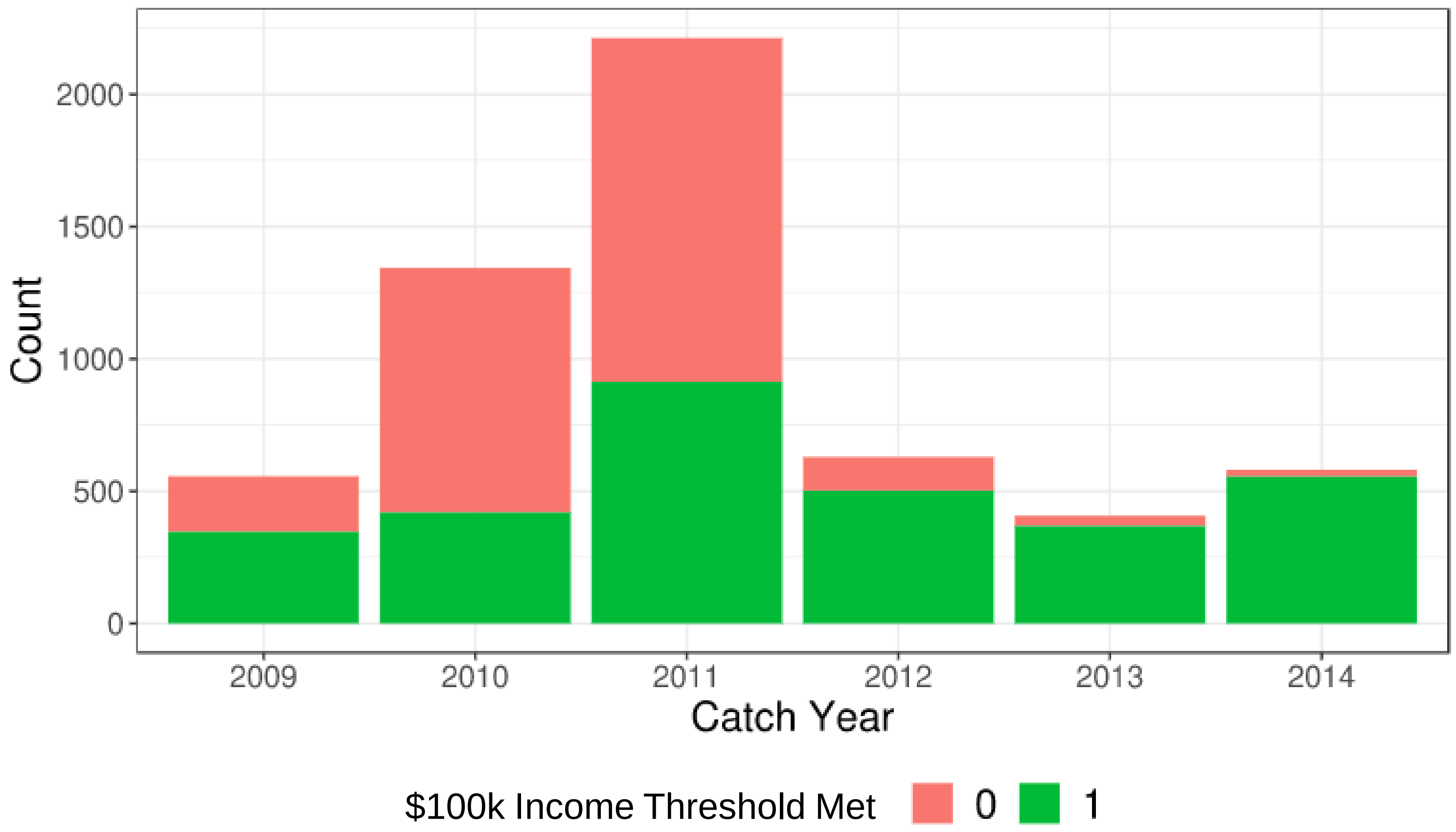


Source: MITRE analysis of CDW data

Control Variable Definitions

- We define taxpayers as ghosts if they did not appear in any of the following inventories:
  - IRMF database containing third-party reported forms,
  - IRTF database containing voluntary reported income tax return forms,
  - IMF database containing records of any activity applied to a taxpayers account,
  - CCNIP database identifying nonfiling taxpayers, and
  - Examination database containing any audit-related interactions with taxpayers.
- Census region of residence was determined from the state derived from the taxpayer's address line or zip code, listed on third-party forms. If census region of residence was not present, region was set to "None".
- Self-employment income is restricted to the types of self-employment income required to be reported to the IRS by third parties: barter income, crop insurance, attorney fees, fishing income, medical payments, non-employee compensation, and patronage income.
- Investment income includes income from distribution shares (Schedule K1), dividends (Schedule 1099-DIV), interest income (Schedule 1099-INT), and passive income (Schedule K1).
- Retirement income includes pension and social security payments.
- Broker transaction income is defined as income from mediating the sale or purchase of property, services, or investments (Schedule 1099-B).
- Other income is defined as income reported on Schedule 1099-MISC, real estate and rental income, lottery income, and business income.

$100k Income Threshold Not Enforced in Baseline Year



Source: MITRE analysis of CDW data

Silver Lining: Estimating the Compliance Response to Declining Audit Coverage

IRS-TPC Research Conference on Tax Administration

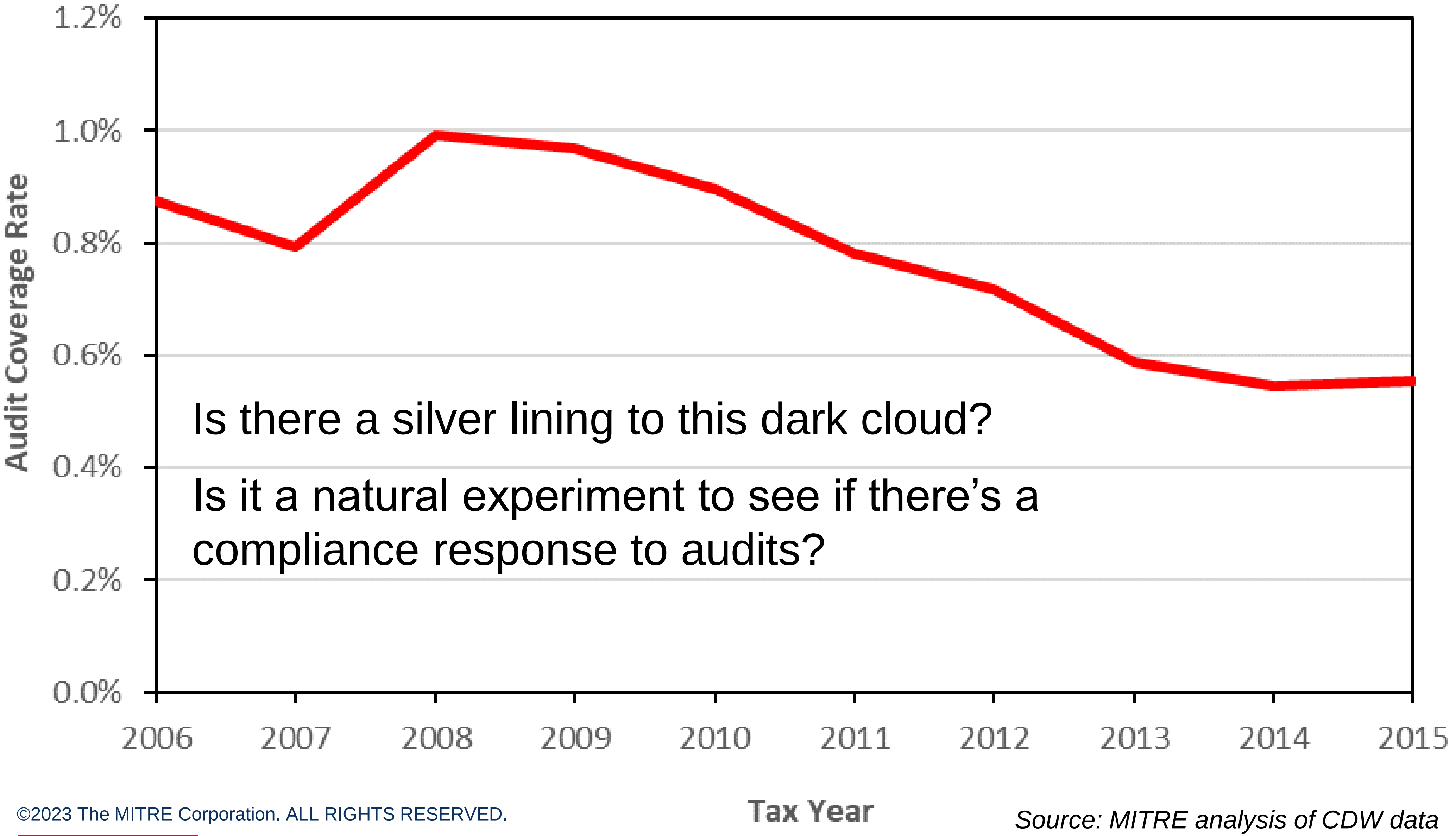
June 22, 2023

Alan Plumley and Daniel Rodriguez (IRS),
Jess Grana and Alexander McGlothlin (The MITRE Corporation)

Research, Applied Analytics & Statistics
KNOWLEDGE DEVELOPMENT & APPLICATION

SOLVING PROBLEMS FOR A SAFER WORLD™

Problem: Audit Rate Steadily Declined



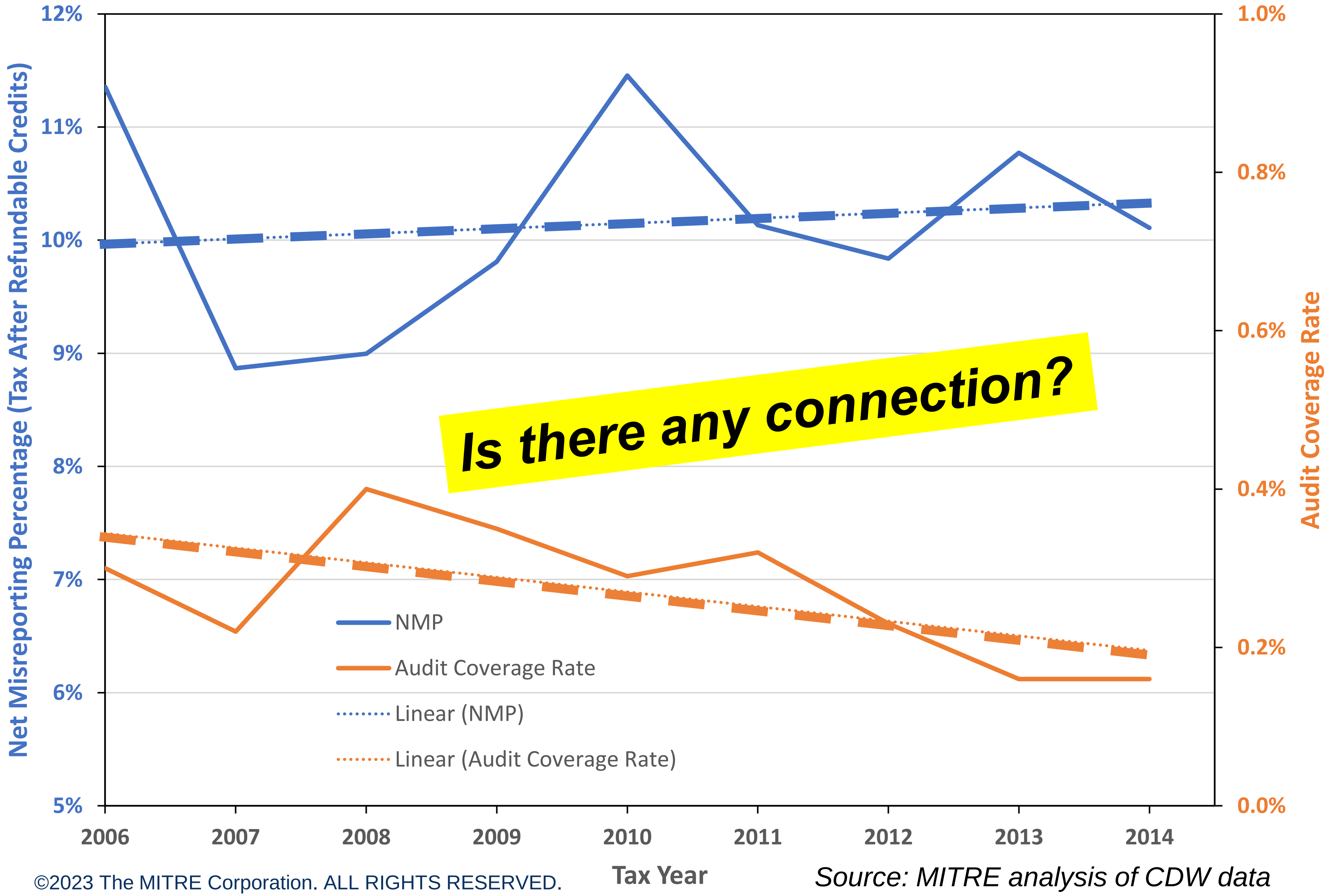
Is there a silver lining to this dark cloud?

Is it a natural experiment to see if there’s a compliance response to audits?

Source: MITRE analysis of CDW data

Largest Category (56%) of Form 1040s

Total Positive Income < $200K; no Business Income/Expenses or EITC



An upward trend in **non**compliance

A downward trend in audit rate

Source: MITRE analysis of CDW data

Maybe… Maybe not

- Not clear for many **other segments** of the population
- Taxpayers likely do not react to (or even know about) contemporaneous trends in audit coverage; their **perceptions may form over time**.
- **Correlation ≠ causation**
  - Other IRS actions?
  - Tax policy changes?
  - Societal trends?

Related Research

- **Definitions:**
  - **Specific** indirect effect: Effect of audit on the *audited* taxpayer's future compliance
  - **General** indirect effect: Effect of audit on *unaudited* taxpayers' future compliance
- **Different levels of "general" indirect effect:**

Demonstration of a "General" Indirect Effect
Demonstrate that a certain type of audit affects the compliance behavior of unaudited taxpayers
Evaluate a given subpopulation (e.g., EITC claimants)
Restrict to a defined network (mechanism)
• ***Tax preparer networks*:** Boning *et al.* (2020); Bohne and Nimczik (2018); Furlong, *et al.* (2021) • ***Supply chain networks*:** Pomeranz (2015) • ***Geographic networks*:** Chetty *et al.* (2013); Drago, Mengel, and Traxler (2020); Alstadsæter, Kopczuk, and Telle (2019); Meiselman (2018); Perez-Truglia and Troiano (2018)

Related Research

- **Definitions:**
  - **Specific** indirect effect: Effect of audit on the *audited* taxpayer's future compliance
  - **General** indirect effect: Effect of audit on *unaudited* taxpayers' future compliance
- **Different levels of "general" indirect effect:**

Demonstration of a "General" Indirect Effect	"Comprehensive" Indirect Effect
Demonstrate that a certain type of audit affects the compliance behavior of unaudited taxpayers	Estimate the overall effect of audit rates on the general population
Evaluate a given subpopulation (e.g., EITC claimants)	Evaluate effects across the taxpayer population
Restrict to a defined network (mechanism)	Agnostic to mechanisms
• ***Tax preparer networks*:** Boning *et al.* (2020); Bohne and Nimczik (2018); Furlong, *et al.* (2021) • ***Supply chain networks*:** Pomeranz (2015) • ***Geographic networks*:** Chetty *et al.* (2013); Drago, Mengel, and Traxler (2020); Alstadsæter, Kopczuk, and Telle (2019); Meiselman (2018); Perez-Truglia and Troiano (2018)	• ***State panel data*:** Dubin, Graetz and Wilde (1990); Plumley (1996); Dubin (2007) • ***Zip code panel data*:** Dubin and Wilde (1988); Grana *et al.* (2022) • ***Microdata (e.g., TCMP)*:** Tauchen, Witte, and Beron (1993); Hoopes, Mescall and Pitman (2012)

This Paper

- **Purpose:** Isolate the indirect effects of audits on the compliance of people not audited (comprehensive indirect effect)
- **Data:** All NRP returns
  - Sample of audits representative of individual taxpayer population
  - Tax Years 2006-2014
- **Method:** Apply econometric techniques to the micro NRP data
  - $ Misreported = *ƒ*(True $, **Audit Rate**, other factors)
  - Misreported and True amounts from the **audit**
  - **Audit rate** is **average** for the return **category**
  - Use **lagged** audit rates due to delay in taxpayer knowledge of IRS enforcement

Noncompliance Measure (Dependent Variable)

- **Net Misreported Amount (NMA)** (in taxpayer's favor)
  - Can be derived for any line item or group of line items
  - For income and tax line items:

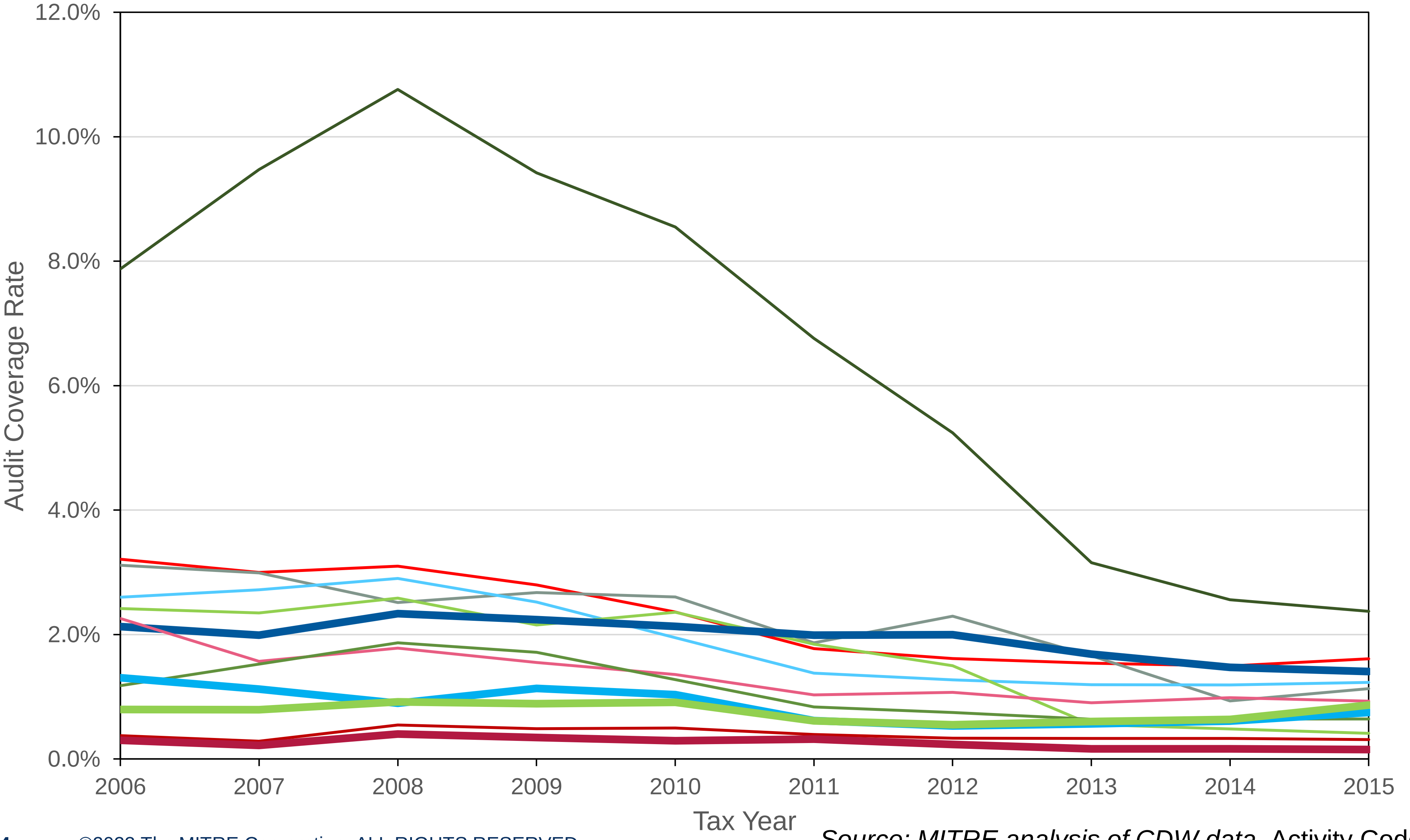
    NMA = $ should have reported - $ reported
  - For offsets to income or to tax:

    NMA = $ reported - $ should have reported

- **Baseline NMA** for Tax After Refundable Credits (TARC)

- **"Visibility Group" NMA** on subsets of line items by visibility of income/offsets

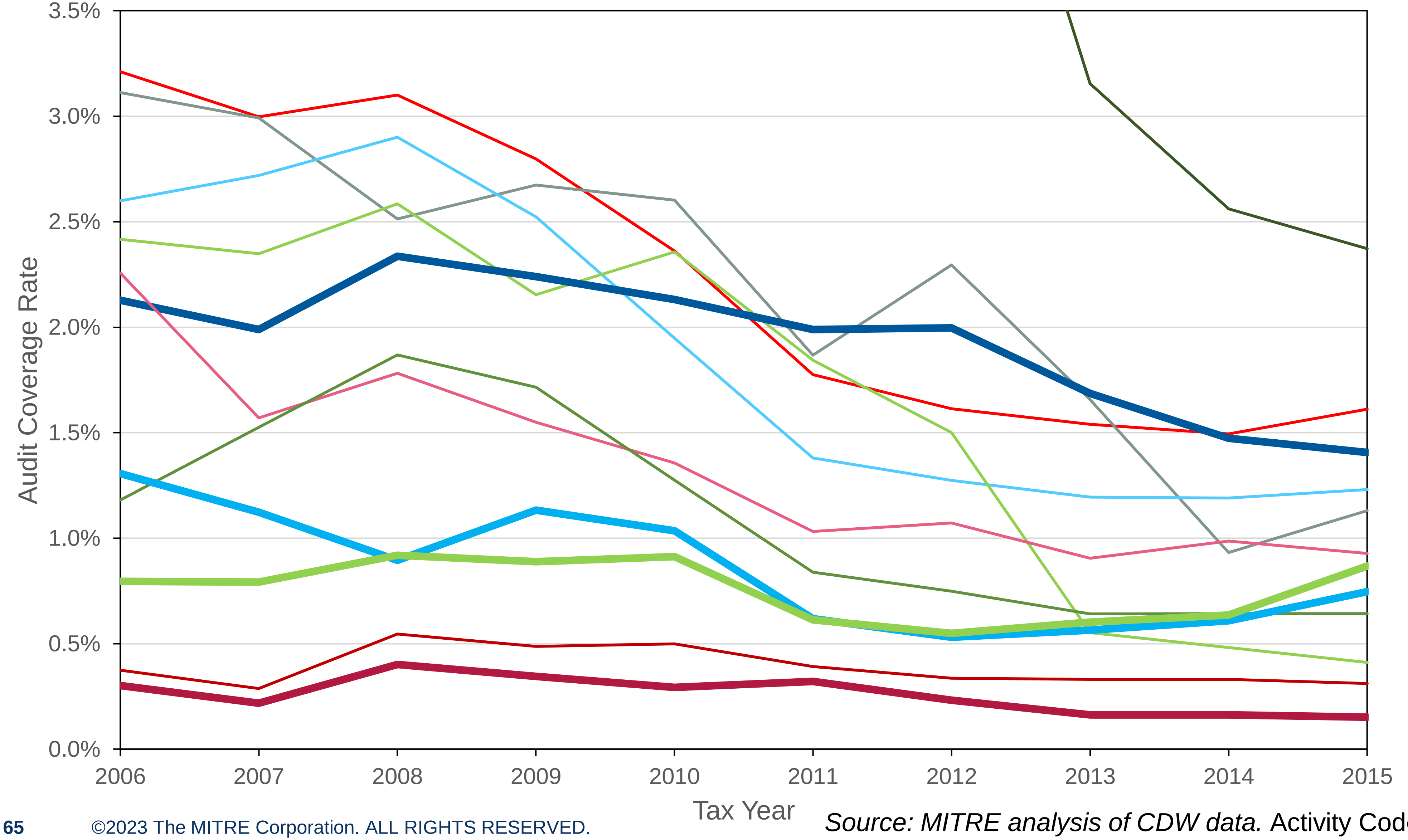
Audit Rates by Return Category (Activity Code)



Activity Code	% of Population
281	0.3%
277	0.5%
280	1.0%
276	0.6%
279	2.4%
270	**17.0%**
271	1.2%
275	2.2%
273	**10.8%**
274	**7.4%**
278	1.0%
272	**55.7%**
	100.0%

Source: MITRE analysis of CDW data. Activity Codes including more than 5% of the population highlighted with bold lines

Audit Rates by Activity Code (All but 281)



Activity Code	% of Population
281	0.3%
277	0.5%
280	1.0%
276	0.6%
279	2.4%
270	**17.0%**
271	1.2%
275	2.2%
273	**10.8%**
274	**7.4%**
278	1.0%
272	**55.7%**
	100.0%

Source: MITRE analysis of CDW data. Activity Codes including more than 5% of the population highlighted with bold lines

Baseline Specification: NMA for TARC

For taxpayer *i* in Activity Code *g* and Tax Year *t:*

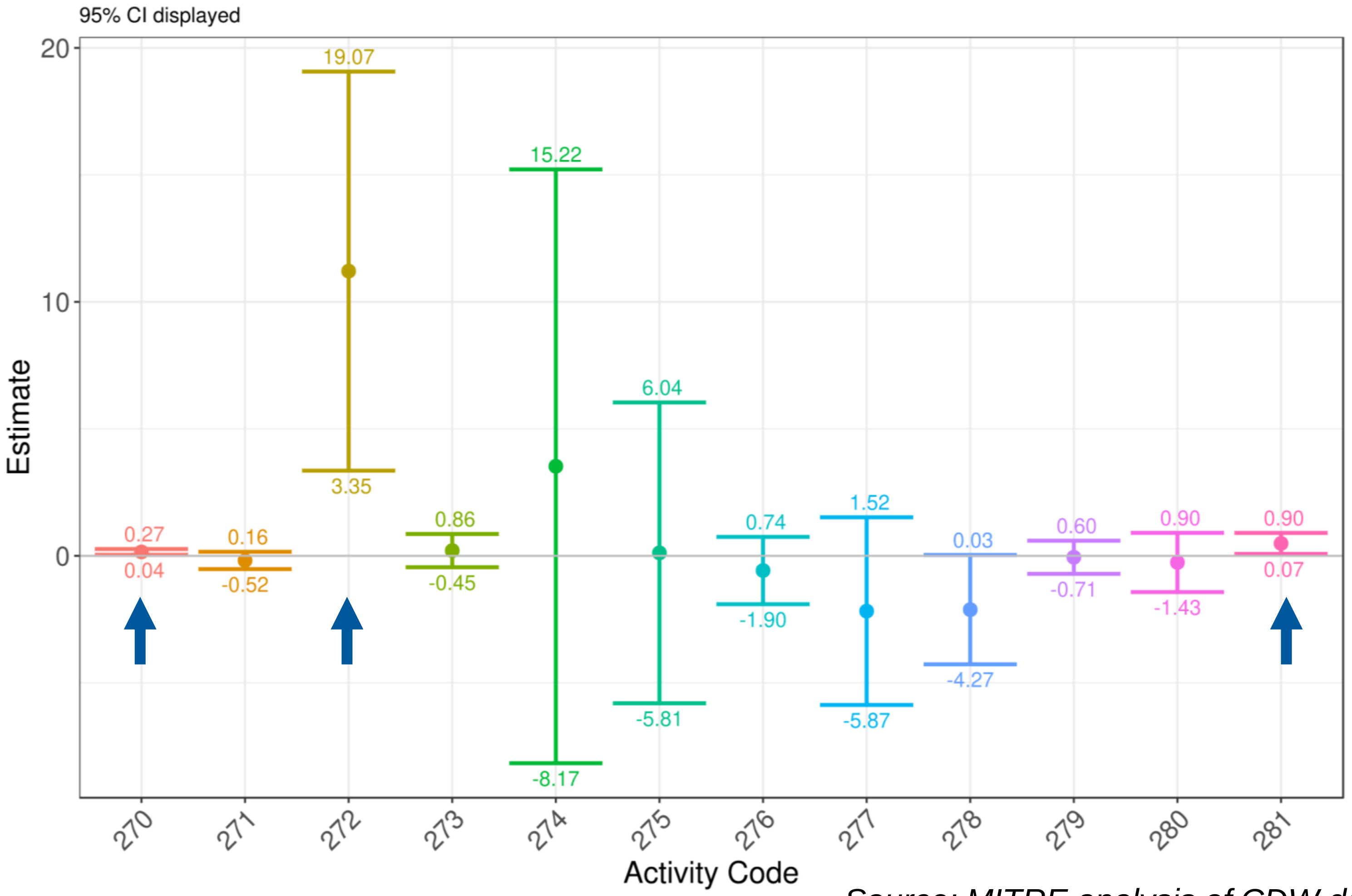
$$\log(NMA\ for\ TARC_{it} + 1)$$
$$= \beta_0 + \boldsymbol{\beta_1 Audit\ Rate_{g,t-2}} + \beta_2 Correct\ Amount\ of\ TARC_{it}$$
$$+ \beta Taxpayer\ Controls_{it} + \alpha Tax\ Year_t + \delta Activity\ Code_g + \varepsilon_{it}$$

β_1	Statistical Significance Level	Taxpayer Controls	
		Positive and Statistically Significant	Negative and Statistically Significant
+0.094	1%	Correct TARC, exemptions, had wages, itemized	Claimed Child Tax Credit (CTC), deducted mortgage interest, over 65, married, filed electronically (ELF)

Unexpected positive effect of audit rate: Perhaps certain subpopulations or noncompliance on certain line items are more sensitive to audit rates…

Subsample Analysis by Activity Code

Effect of Lagged Audit Rate on **NMA for TARC**



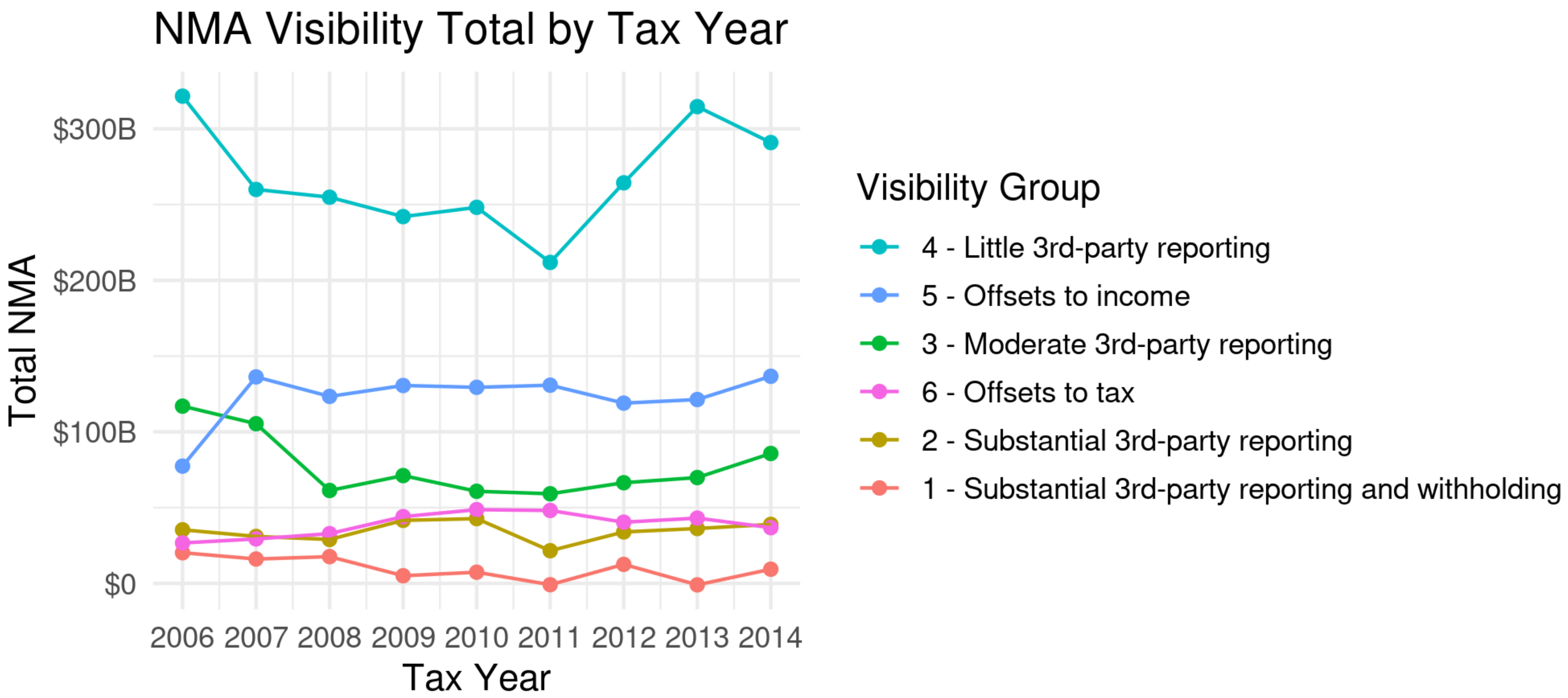
Source: MITRE analysis of CDW data

Activity code defined by:

- Income range
- EITC claiming
- Business, Non-business

> ➢ Unexpected positive effect of audit rates on 270, 272 and 281
>
> ➢ Negative effect on 278 is significant at 10% level

What About Different Line Items?



Source: MITRE analysis of CDW data

Dependent Variable: NMA by Visibility Group (VG)

$$\log(NMA\ for\ VG_{it} + 1)$$
$$= \beta_0 + \boldsymbol{\beta_1 Audit\ Rate_{g,t-2}} + \beta_2 Correct\ Amount\ for\ VG_{it} + \boldsymbol{\beta} Taxpayer\ Controls_{it}$$
$$+ \alpha Tax\ Year_t + \delta Activity\ Code_g + \varepsilon_{it}$$

Visibility Group	β_1	Taxpayer Controls	
		Positive and Significant	Negative and Significant
1	**-0.036 *** **	Had wages, married, ELF	Correct amount, exemptions, CTC, itemized, >65, paid prep
2	0.001	Correct amount, exemptions, had wages, itemized, >65, married	CTC, paid prep, ELF
3	0.003 **(-0.030 * w/o TY FE)**	Correct amount, mortgage, >65, paid prep, married	Exemptions, had wages, CTC, itemized
4	**+0.057 ** **	Correct amount, exemptions, mortgage, paid prep, married	Had wages, CTC, itemized, >65, ELF
5	-0.040	Exemptions, had wages, itemized	Correct amount, CTC, mortgage, >65, paid prep, ELF, married
6	0.021	Exemptions	Correct amount, itemized, mortgage, >65, married

*** 1%, ** 5%, * 10% level of significance

Sensitivity Analyses

- Do taxpayers respond to a ***3-year lag*** of audit rate? Or ***4***?
    - NMA for TARC: Unexpected positive effect of 2-year lag **reverses** when using 4-year lag (not significant)
- Do only ***certain taxpayers*** adjust ***certain line items***?
    - NMA for lower visibility line items: Expected **negative** effect for higher income taxpayers
- Do taxpayers respond to a more ***aggregate audit rate***, such as across similar Activity Codes?
    - NMA for TARC: Unexpected positive effect for some Activity Codes **reverses** when using more aggregated audit rates
- Do taxpayers respond to ***spending on audits*** rather than rates?
    - NMA for TARC: Expected negative effect for **more** activity codes

Discussion

- Very few estimates of how enforcement affects <u>overall</u> compliance of the general population
- **Findings**:
  - Misreporting on high visibility income (wage and salaries) drops by 3.6 to 6.1 percent with a one percentage point increase in audit rates.
  - For other line items, indirect effect detected for only certain taxpayers. Some unexpected positive effects reverse in sensitivity analyses.
  - Results are mixed on misreporting by taxpayers earning above $200,000 who earn business income – *but difficult to validate true income at the high end*.
- **Next steps**:
  - Disaggregate some Visibility Groups, convert estimates to dollar values, econometric extensions

COMMENTS ON “CHANGES TO VOLUNTARY COMPLIANCE FOLLOWING RANDOM TAXPAYER AUDITS”, “THE LONG-TERM IMPACT OF AUDITS ON NONFILERS’ TAX COMPLIANCE”, AND “SILVER LINING: ESTIMATING THE COMPLIANCE RESPONSE TO DECLINING AUDIT COVERAGE”

William Boning, U.S. Department of the Treasury

IRS-TPC Conference, June 2023

Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the Department of the Treasury.

NEAT EXPERIMENTS, SOME SHARED CHALLENGES

- Not always clear what the bottom line is
  - On the road from “what we tried” to “what you need to know”
- Statistical power
  - We would expect similar results for similar subgroups
    - Except for noise from low power
  - Are tests over-rejecting? Maybe bootstrap SE
  - Consider Buonferroni correction when running several tests

And failing to reject the null of no effect doesn’t mean the effect is zero.

COMMENTS ON BESNEK AND PARTINGTON I

First, a clarification question about selection:

1. The REP is randomly selected
2. Then returns are profiled and income is matched
3. Then issues are reviewed
4. Finally, some returns are audited

Is the analysis sample the whole REP (1) or only audited returns (4)?

- (1) is more comparable to controls and the overall population

COMMENTS ON BESNEK AND PARTINGTON II

Focus on novel contributions:

- Specific deterrence for small corporations
- Questions the Australian tax system is especially good for answering
  - E.g. Is compliance even more closely related to information returns when tax returns are pre-filled?

COMMENTS ON BESNEK AND PARTINGTON III

Simple is powerful. Use the whole sample - don’t split by year or audit outcome.

- Differences across years are probably noise
- The cost-benefit analysis for the audits depends on the overall average.
- The split by audit outcomes mostly confirms that the audits drive the differences
- Instead: most common kinds of non-compliance detected

COMMENTS ON BESNEK AND PARTINGTON IV

Estimate effects by years since (or before) audit

$$y_{it} = \sum_j a_j D_j + \beta_j D_j D_{treated} + \sum_t \sigma_t D_t + e_{it}$$

where j is years since (or before) selection for audit or as a control and t is calendar years.

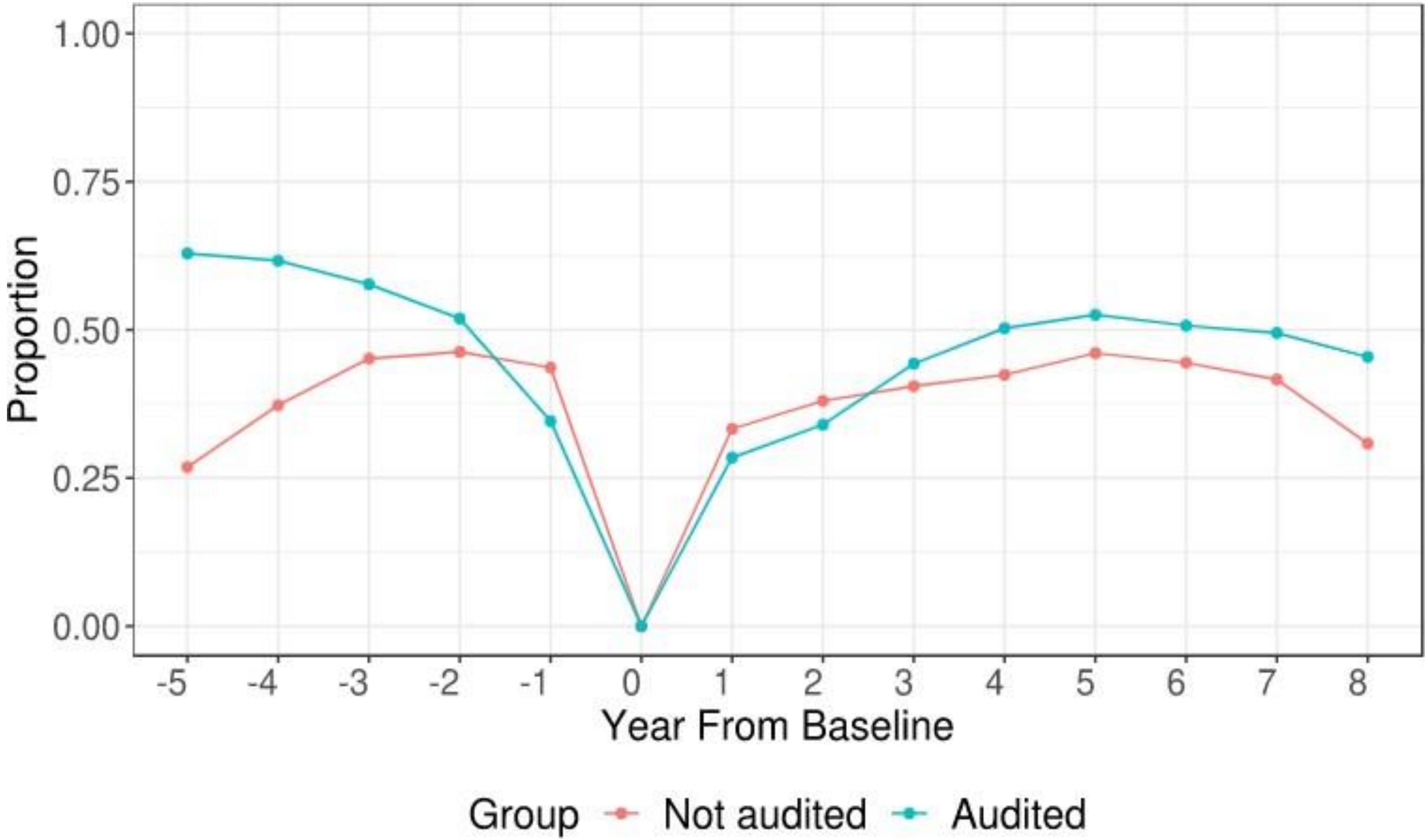
- Years before audit are a placebo test
- Trends over time stand out on a graph
  - How quickly does the effect fade?
- What's the total (present-discounted) return summed over all the post-audit years?

COMMENTS ON LINDSAY, GRANA, & PLUMLEY NONFILERS I

Nice clear question and bottom line. Can you expand on your contributions relative to Datta et al. (2015)?

- Audits are a more intense treatment than automated substitutes for returns.
- Are your methods an improvement?
- Which population contributes more to the filing part of the tax gap?

IDENTIFICATION CONCERNS



IDENTIFICATION SUGGESTIONS

This is a good context for matching

1. Audits could only have been selected based on things you observe.
2. Use the baseline and year-prior characteristics to predict audit within your treatment and control groups.
3. Use the generated propensity scores for propensity score matching or inverse probability weighting.
4. Check that baseline characteristics and prior year trends are similar for control and treatment.

COMMENTS ON PGRM SILVER LINING I

Lean on theory for guidance.

- It's all about perceived p of detection
- Competing hypotheses about how perceptions change when audit rates change:
  1. Total ignorance
  2. Hazy, lagged idea about the change
  3. Perfect information
- Doesn't p of detection for wages and salaries depend on document matching rather than audit rates?

COMMENTS ON PGRM SILVER LINING II

The big picture under the hazy perceptions hypothesis:

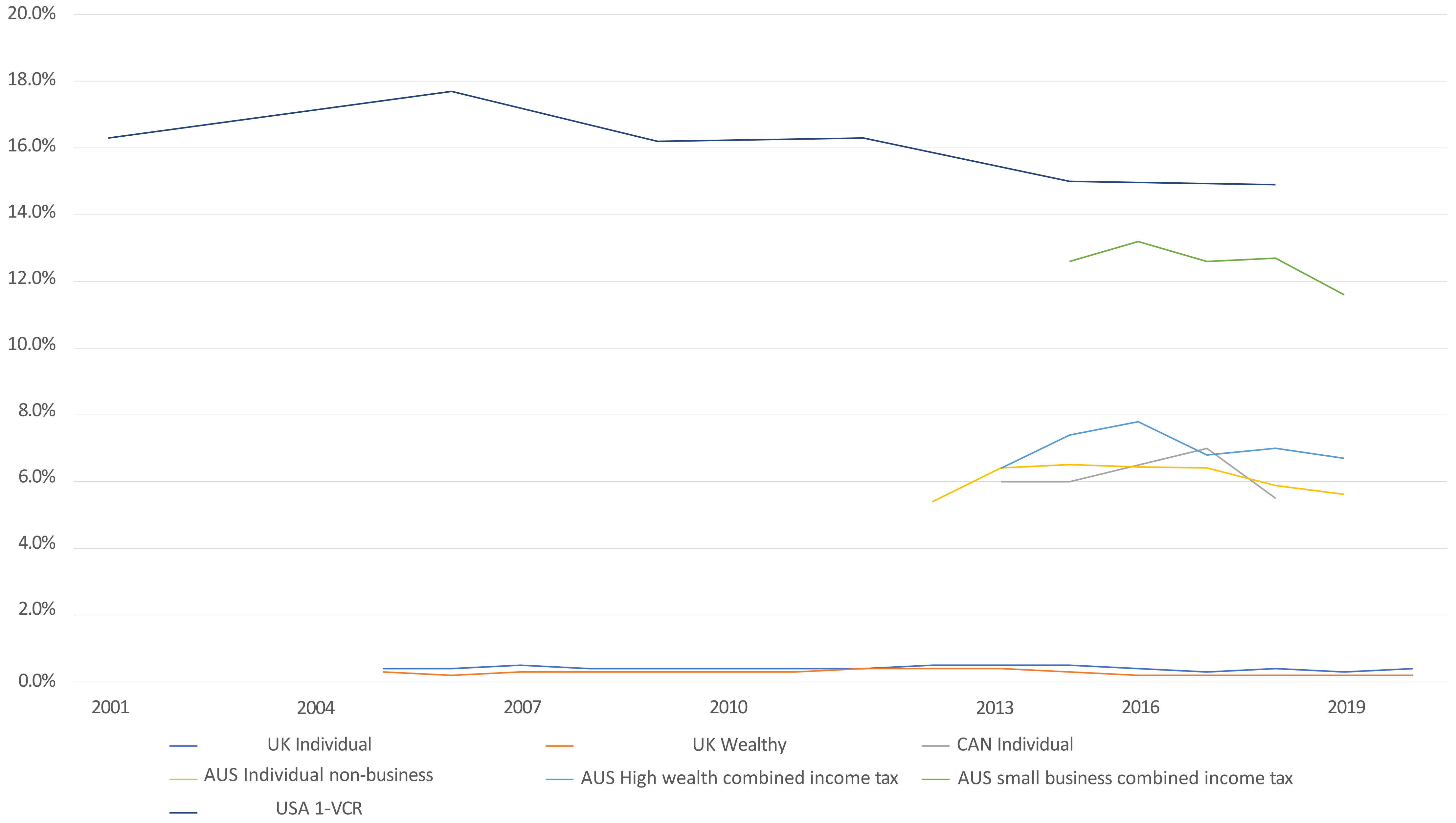
- Audit rates have fallen a lot, especially at the top
- Aggregate noncompliance (NMA) has been pretty much flat
  - But we would expect aggregates grow over time
  - Has noncompliance as a percentage (NMP) fallen despite lower audit rates?

COMMENTS ON PGRM SILVER LINING III

Use the bigger decline in audit rates for high income:

- Split into fewer, more meaningful groups than activity codes
  - For maximum power: high income vs. the rest
  - Another option: business/EITC/high income/other
- What is NMP over time for each grouping?

TAX GAPS TRENDING FLAT OR DOWN AROUND THE WORLD



WHERE DO WE GO FROM HERE?

Directions for future research:

- What can we learn by comparing countries?
  - How much do technologies like e-filing and document matching matter?
- Surveys or lab experiments on the non-monetary costs (hassle/psychological) costs of being audited
- Surveys on perceptions of audit rates, perceived changes over time, perceived differences across groups