Indirect Deterrence Effects from IRS Filing and Payment Compliance Programs

Brett Collins, Corbin Miller, Mark Payne, Sean Roh, Yan Sun, Alex Turk, and Chris Wilson
(IRS Research, Applied Analytics, and Statistics)[1]

Abstract

This study examines how IRS enforcement strategies—collection notices, return-filing reminders, and field visits—indirectly influence previously compliant taxpayers and sustain their compliance. Using a two-stage multinomial logistic model, we estimate the causal spillover effects while addressing potential endogeneity through IRS workforce changes as an exogenous instrument. To quantify these effects, we leverage the Social Connectedness Index (SCI) to measure exposure to enforcement in socially linked areas, capturing how enforcement actions influence taxpayer behavior beyond directly treated individuals. We find that a 10% increase in enforcement reduces newly accrued delinquent amounts by 16% ($3.2 billion) for collection notices and 7% ($1.3 billion) for return-filing reminders. Letters, with their broader reach and frequency, yield stronger indirect effects than field visits, which are more targeted but resource-intensive and address fewer high-debt cases. These findings highlight the tradeoff between broad, frequent actions fostering indirect compliance and intensive, direct interventions addressing severe delinquencies.

I. Introduction

The Internal Revenue Service (IRS) plays a central role in maintaining the integrity of the U.S. tax system, which relies on voluntary compliance. In Fiscal Year 2023, the IRS processed over 271 million federal tax returns and supplemental documents (Internal Revenue Service 2024a). While most taxpayers meet their filing and payment obligations without direct intervention, ensuring sustained compliance remains a challenge. In Tax Year 2022, approximately 85% of total tax liabilities were paid voluntarily and on time, yet a substantial portion remained unpaid, contributing to a tax gap of $696 billion (Internal Revenue Service 2024b)[2]. Even after enforcement efforts, a significant share of these unpaid liabilities—$606 billion—remains uncollected, underscoring the persistent difficulty of achieving full compliance and highlighting

[1] The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect the views or the official positions of the U.S. Department of the Treasury or the Internal Revenue Service. All results have been reviewed to ensure that no confidential information is disclosed.

[2] The tax gap is the difference between the total true tax liability owed by taxpayers for a given tax year and the amount that is paid voluntarily and on time. It consists of three components: (1) Non-filing—tax not paid on time by those who do not file required returns, (2) Underreporting—tax that is understated on timely filed returns, and (3) Underpayment—tax that is reported but not paid on time.

the critical role of enforcement strategies in influencing taxpayer behavior (Internal Revenue Service 2024b).

The ability of the IRS to close this tax gap through enforcement, however, has been increasingly constrained by limited resources. Between 2010 and 2019, the agency's enforcement budget declined by more than 28% in real terms, while its responsibilities expanded due to legislative changes and increased administrative burdens. Over the same period, the number of full-time equivalent (FTE) employees dedicated to enforcement fell by 34%, from 50,400 to 33,484, reducing the IRS's capacity to conduct audits and pursue delinquent taxpayers (Internal Revenue Service 2011, 2020). These constraints necessitate a more strategic allocation of enforcement resources to not only recover unpaid taxes but also maximize compliance through deterrence effects. A key question, therefore, is how enforcement efforts—particularly those related to filing and payment compliance—affect taxpayer behavior, both directly and indirectly.

Prior research distinguishes between direct enforcement effects, which apply to taxpayers who receive an enforcement action, and indirect effects, where enforcement influences individuals who were not directly contacted but adjust their behavior based on perceived risk or awareness of IRS activities. A growing body of literature highlights the importance of these spillover effects, suggesting that enforcement actions can shape compliance norms within communities and networks. For instance, Boning et al. (2019) demonstrate that IRS field visits not only increase compliance among targeted firms but also among businesses connected through the same tax preparer network. While these studies underscore the role of social networks in amplifying enforcement effects, much of the literature focuses on non-compliant taxpayers. The extent to which IRS enforcement actions reinforce compliance among previously compliant taxpayers remains an open and underexplored question.

This study seeks to fill that gap by investigating the indirect deterrence effects of IRS enforcement on taxpayers who were compliant in the previous year but may become delinquent in the current year. Specifically, we examine how IRS enforcement actions function as a preventative mechanism, sustaining voluntary compliance among historically compliant taxpayers. Our analysis focuses on three key enforcement actions related to filing and payment compliance: Automated Collection System (ACS) notices, which are mailed reminders sent to taxpayers with outstanding balances; CP59 notices, which target non-filers, requesting submission of overdue returns; and field collection visits, where IRS revenue officers conduct in-person interventions to address persistent delinquencies. These enforcement strategies vary in their intensity and reach. ACS and CP59 notices are correspondence-based enforcement tools, allowing the IRS to contact a broad population of taxpayers at a relatively low cost. Field collection visits, in contrast, are resource-intensive and geographically localized, with revenue officers directly engaging delinquent taxpayers. While these actions are primarily designed to address existing noncompliance, their visibility within communities may influence taxpayers who have not yet fallen into delinquency, reinforcing the perceived risk of noncompliance and encouraging continued compliance.

To assess these indirect effects, we leverage a two-stage least squares (2SLS) regression model with instrumental variables, using fluctuations in IRS enforcement resources from 2011 to 2019 as a natural experiment. These fluctuations provide a natural experiment to help isolate the causal impact of enforcement actions from potential confounding factors. Additionally, we employ the Facebook Social Connectedness Index (SCI) to capture how enforcement awareness spreads through social networks, rather than relying solely on geographic proximity for a proxy of enforcement exposure. This approach allows us to quantify how enforcement actions propagate compliance effects beyond directly treated individuals and across socially connected communities.

Our findings reveal that IRS enforcement actions have significant indirect deterrence effects on previously compliant taxpayers. ACS and CP59 notices, in particular, generate measurable spillover effects due to their broad reach and visibility. We estimate that a 10% increase in ACS notices reduces newly accrued delinquent balances among previously compliant taxpayers by 16%, highlighting the substantial compliance benefits of scalable, correspondence-based enforcement. Moreover, these effects are amplified in regions with higher social connectedness, suggesting that enforcement actions influence taxpayer behavior through both direct treatment and social spillovers.

This study makes several key contributions to the tax compliance literature. First, it broadens the scope of enforcement research by demonstrating how compliance interventions can sustain voluntary compliance rather than merely rectifying noncompliance. Second, by incorporating network-based measures of enforcement exposure, our analysis offers a more comprehensive understanding of how taxpayers perceive enforcement risk. Third, the utilization of instrumental variables and natural fluctuations in enforcement resources allows us to identify causal relationships between enforcement actions and taxpayer behavior. Furthermore, our results imply that conventional estimates of the direct effects of low-cost, frequent enforcement actions may substantially understate their total impact by neglecting spillovers to compliant taxpayers via social networks. Recognizing the influence of enforcement salience among social contacts could enhance the design and effectiveness of compliance programs.

The remainder of the paper is organized as follows. Section II reviews the institutional background of IRS enforcement mechanisms and the literature on tax compliance spillovers. Section III details the data sources and empirical methodology, including our instrumental variables and network-based enforcement measures. Section IV presents the primary empirical findings, quantifying indirect enforcement effects. Finally, Section V discusses policy implications and concludes with suggestions for future research.

II. Background

A. Program Trends

The IRS enforcement budget declined by more than 28% in real (inflation-adjusted) terms from 2010 to 2019, as illustrated in Figure 1. This budget contraction occurred alongside an increasing number of tax returns to process and growing administrative responsibilities related to new legislation and compliance issues. The smooth downward trend in budget figures masks the substantial challenges faced during this period, including rising instances of identity theft, multiple federal government shutdowns—most notably the longest in U.S. history during 2018-2019—and the need to adapt to significant legislative changes such as the Tax Cuts and Jobs Act of 2018.

To manage its expanding responsibilities amid constrained resources, the IRS had to redistribute its workforce across various programs, leading to cutbacks in several filing and payment compliance initiatives. Between 2010 and 2019, the number of FTE employees at the IRS decreased by 34%, from 50,400 to 33,484 (Internal Revenue Service 2011, 2020). Figure 1 illustrates the relationship between enforcement budget trends and FTE allocations across different enforcement activities. While the overall trend shows a decline in staffing, there are notable shifts in how FTEs were allocated to different enforcement types. These shifts reflect administrative decisions that were shaped by resource constraints and broader enforcement priorities rather than direct responses to individual taxpayer compliance behavior.

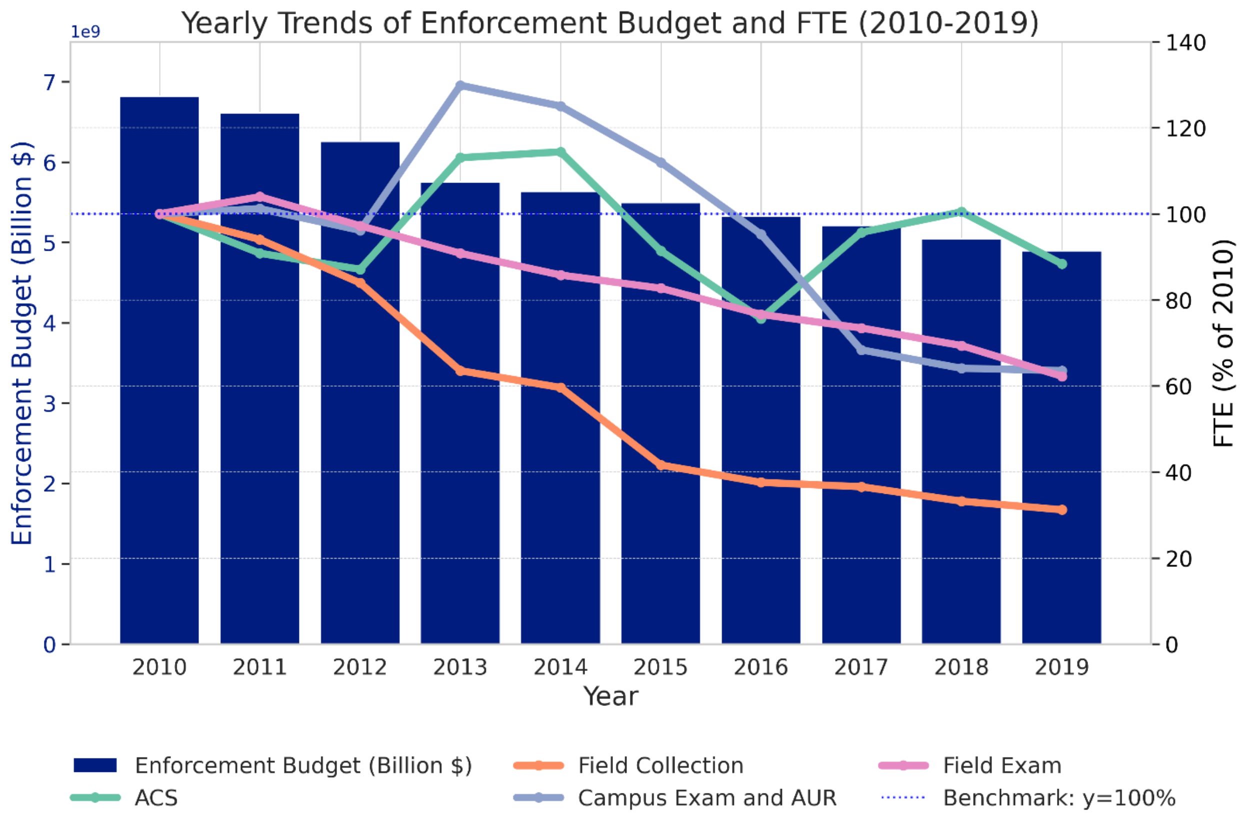


Figure 1. IRS Enforcement Budget and FTEs

Note: The bar chart represents the annual enforcement budget from 2010 to 2019. The line graphs show the annual FTE positions for various enforcement programs, including the ACS, Campus Examinations, Automatic Underreporter (AUR), Field Collection visits, and Field Examinations. All line graphs are normalized to 2010 as the base year to allow for consistent trend comparison across different FTE scales. *SOURCE: SOI Data Book, Table 30, inflation adjustment calculated with Bureau of Labor Statistics CPI-U consumer price index*

These heterogeneous changes in the allocation of FTE positions subsequently affected the number of enforcement actions—such as ACS letters, CP59 notices, and field visits—conducted under each type of program. This variation in enforcement intensity, driven by exogenous shifts in FTEs, creates a natural experiment that allows us to tease out the causal effect of enforcement activities on taxpayer compliance. Such insights are particularly valuable for optimizing resource allocation, especially given the increased funding for enhanced enforcement efforts following the Inflation Reduction Act of 2022.

Reductions in appropriations and FTEs coincided with increases in the numbers of taxpayers in a delinquent status. As shown in Figure 2, which tracks key trends for individual Form 1040 taxpayers over our study period by using 2010 as a base year, the number of taxpayers with unpaid assessments (UA) and those identified as non-filers through the Case

Creation Nonfiler Identification Process increased. This rise in noncompliant taxpayers contrasts with a decline in enforcement actions, as reflected by the reduced issuance of delinquent return notices (CP59), selected ACS letters, and field collection assignments. Within these overall declines, there is considerable variation in both the timing and intensity of these enforcement programs, which enhances the ability of our models to isolate each program's impact on compliance.

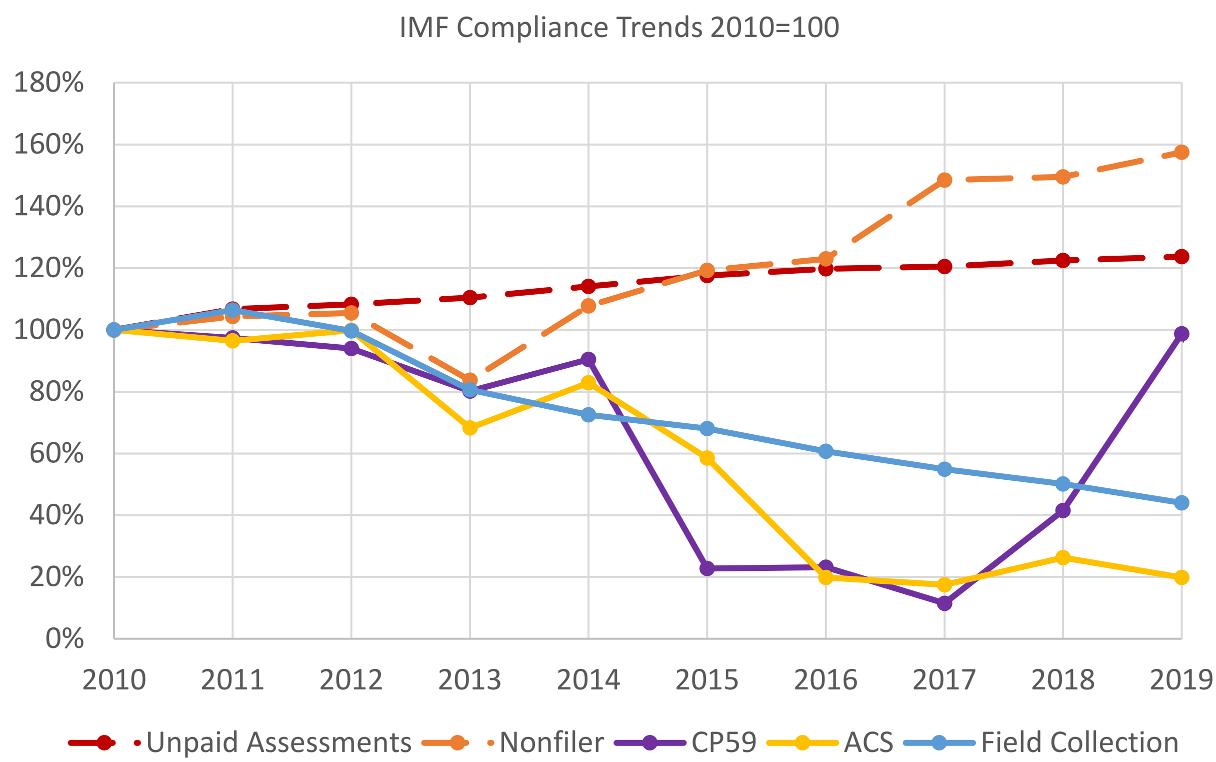


Figure 2. Compliance Trends

SOURCE: Compliance Data Warehouse, counts of non-compliant taxpayers (dotted lines) and IRS treatments (solid lines) indexed to 2010

B. Program Operation

The IRS collection process begins with a return matching system that cross-references taxpayer-reported income against third-party sources, such as employer-reported wages and financial institution filings. If a taxpayer fails to file a return, this matching process cannot proceed, and the case is placed into the delinquent return inventory. Regardless of filing status, taxpayers with unpaid assessments enter the collection process (Internal Revenue Service 2024b). Upon entering the collection process, taxpayers typically receive a balance due notice, informing them of their outstanding liability and providing instructions for resolution (Figure 3). Non-filers may receive a CP59 notice, which notifies taxpayers that the IRS has no record of a filed tax return for a given year. Recipients are advised to submit their return immediately or provide justification for not filing. If the taxpayer does not respond or resolve their liability after these initial notices, the IRS may escalate enforcement efforts by opening a Taxpayer Delinquent Account (TDA), which may trigger further collection actions.

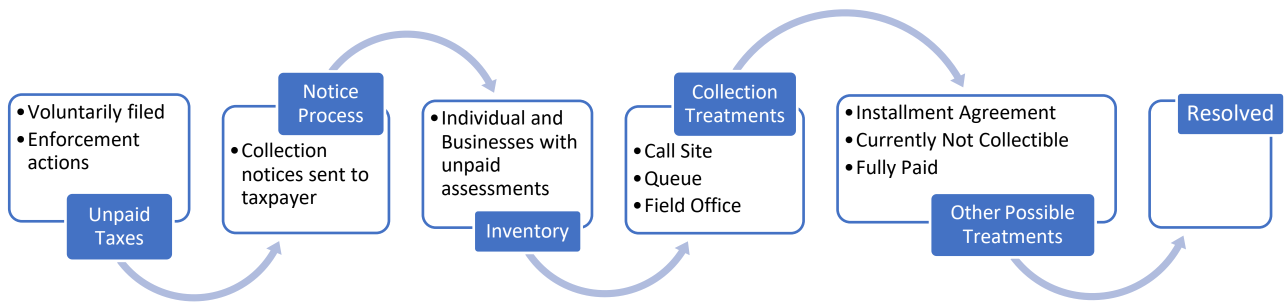


Figure 3. IRS Collection Process

Cases are assigned to different IRS collection programs based on factors such as the size of the unpaid balance, the complexity of the taxpayer's financial situation, and available enforcement resources. The IRS employs both automated and in-person enforcement mechanisms. The Automated Collection System (ACS), which operates from IRS campus facilities, issues a series of notices to prompt taxpayer compliance. Among these, LT11 (Final Notice of Intent to Levy and Your Notice of a Right to a Hearing) serves as the final warning before the IRS proceeds with asset seizure, informing taxpayers of their right to contest the levy. The LT16 notice is a reminder urging immediate resolution of unpaid balances to prevent potential enforcement actions, such as levies or liens. The LT26 notice is directed at non-filers who have ignored prior IRS communications, demanding that they file their outstanding tax returns (Internal Revenue Service n.d.).

For cases requiring in-person enforcement, Revenue Officers (ROs) conduct field collection visits, typically reserved for high-priority cases involving significant unpaid balances, uncooperative taxpayers, or complex financial circumstances. Field visits allow the IRS to obtain financial disclosures, issue levies, or negotiate installment agreements directly. However, in a recent policy shift, the IRS has largely ended unannounced visits by revenue officers to improve taxpayer safety and reduce confusion. Instead, ROs now initiate contact through an appointment letter (IRS Form 725-B) to schedule meetings. Depending on the taxpayer's response, their case may be reassigned to different enforcement treatments, placed in the IRS "queue" pending further review, or resolved through full payment, installment agreements, or designation as Currently Not Collectible (CNC) status, the equivalent of a write off.

In this study, we focus on ACS notices (LT11, LT16, and LT26) and CP59 notices, as well as field collection visits by ROs. These interventions differ in their implementation: ACS and CP59 notices are remote enforcement mechanisms issued from centralized IRS facilities, whereas field visits involve direct engagement by local IRS offices. Given their higher resource intensity, field visits are much less common than ACS notices. A critical distinction in IRS enforcement is between discretionary enforcement actions, such as ACS notices and field visits, and automatic collection procedures, such as balance due notices. While balance due notices represent a mandatory early-stage enforcement step, they do not constitute discretionary enforcement actions. Unlike ACS notices, which can be intensified or strategically deployed based on IRS priorities, balance due notices are systematically issued to all taxpayers with outstanding balances (Internal Revenue Service 2024a).

Because this study seeks to estimate the causal effects of discretionary enforcement actions, we exclude balance due notices from our treatment variables. These notices lack exogenous variation, making it difficult to disentangle their compliance effects from broader systemic enforcement trends. While the issuance of balance due notices was temporarily disrupted during the COVID-19 pandemic, isolating the compliance effects of this pause from other pandemic-related factors—such as economic stimulus payments and temporary IRS enforcement suspensions—falls beyond the scope of this analysis.

C. Literature Review

The impact of IRS compliance programs can be broadly categorized into direct effects—changes in current and future behavior for taxpayers subject to enforcement—and indirect effects, where non-treated taxpayers adjust their behavior based on perceived IRS enforcement activity. Indirect effects suggest that non-treated taxpayers acquire information about enforcement likelihood through various channels, such as preparer networks, public data on IRS activities, social circles, or news outlets. These effects are particularly relevant for tax administration, given the critical role of voluntary compliance in the U.S. tax system (Bloomquist, 2012; Datta et al., 2015; Boning et al., 2019). Typically, in the literature, the success of IRS compliance programs is evaluated based on observed changes in taxpayer behavior, such as timely filing and payment or the resolution of prior delinquencies.

Most previous studies that estimate indirect effects have focused on specific programs but demonstrate notable indirect effects. For instance, Datta et al. (2019) analyzed the Automated Substitute for Return (ASFR) program, finding that indirect effects increased the likelihood of filing for non-treated taxpayers by up to 27%, surpassing direct effects and persisting over time. Similarly, Turk and Ashley (2002) examined the Notice of Federal Tax Lien (NFTL) program and leveraged a policy change to assess both the direct and indirect effects on delinquent taxpayers' likelihood of resolving their tax debts. Although indirect effects have gained increasing attention, the research remains constrained due to technical complexities and data limitations (Boning et al., 2020). Studies in this area typically rely on three methodological approaches: field experiments, laboratory experiments, and natural experiments.

Field Experiments

Lopez-Luzuriaga and Scartascini (2023) conducted a field experiment in Argentina in 2011 to examine how various interventions affected non-payers' compliance with unpaid property taxes. They compared three types of treatment messages: deterrence, reciprocity, and peer-effect. The study found that a deterrence letter—emphasizing penalties and the likelihood of detection—was the most effective in increasing compliance.

Their model predicted that taxpayers with higher tax morale or risk aversion are more likely to comply, while liquidity constraints pose challenges to compliance. Although the focus was on direct effects, the study reinforced that the perception of penalties and detection probability are key factors in tax compliance, echoing findings from Boning et al. (2019). Furthermore, the dissemination of information on penalties and detection is not restricted to the treated taxpayers but spreads through social networks, suggesting that compliance behaviors may change through indirect channels such as group effects (Bloomquist, 2012) or network effects (Boning et al., 2019).

Boning et al. (2019) conducted a randomized experiment in 2015 to study both direct and indirect effects of IRS enforcement on employer Federal Tax Deposit collections. The study tested the effects of sending letters and conducting in-person visits to at-risk firms. They found that in-person visits had significant and persistent direct effects on tax payments, while letters had smaller effects. The study specifically examined network effects, where information about enforcement activities spread through shared tax preparers. Firms whose tax preparers had other clients receiving in-person visits from IRS Revenue Officers were more likely to remit taxes. The aggregate network effect was larger than the direct effect, producing 1.2 times more revenue. No similar network effect was observed for letters.

Agent-Based Models

Bloomquist (2012) emphasized the importance of indirect effects in tax administration and identified three types of these effects: induced, subsequent period, and group effects. Although Bloomquist considered changes in taxpayer behavior due to prior audits as indirect, we treat these as direct effects. More relevant are group effects, where individuals alter their behavior based on others' experiences, such as those within the same community or workplace. Bloomquist estimated that every $1 detected through audit selection generates $6 to $11.60 in indirect effects. Bloomquist developed an Agent-Based Model (ABM) to quantify these indirect effects, using artificial taxpayer data from Tax Year 2001 to simulate income tax reporting behavior in a small region. The model showed that audit selection strategies incorporating indirect effects—particularly group effects—yielded a greater impact on voluntary compliance, as taxpayers adjusted their reporting behavior based on the audit experiences of their neighbors and coworkers.

Natural Experiments

Using a natural experiment stemming from declining IRS budgets and reduced enforcement activity, Datta et al. (2015) analyzed the direct and indirect impacts of the IRS ASFR program on delinquent tax collections and subsequent compliance. The study also estimated the program's indirect effects on non-filers more broadly. The dataset comprised a random 10% sample of delinquent tax returns from tax years 2007 to 2009, and subsequent returns. The study first calculated the predicted probability of taxpayers being selected for ASFR treatment to capture indirect effects, followed by estimates of revenue collected in subsequent years. Results showed significant direct and indirect impacts on compliance. Treated cases generated $672 to $1,640 in revenue, depending on the model, while untreated cases exhibited indirect effects ranging from $194 to $1,187. The study also identified stronger, longer-lasting indirect effects on filing compliance.

Turk et al. (2016) examined the direct effects of the IRS NFTL program on delinquent tax collections for individuals and businesses. The study used a policy change in NFTL filing thresholds in 2011 as a natural experiment, tracking taxpayer outcomes for two years after cases were transferred from the ACS to the Field Collection Queue. The research found that NFTLs significantly increased the likelihood of reducing outstanding balances, with individual taxpayers' balances falling by 22-23% over one year and 33-35% over two years. Business taxpayers experienced larger reductions, ranging from 38-40% over one year to 60-65% over two years.

III. Data and Methodology

A. Data

The study population consists of individual filers who fully paid and timely filed their return in the prior year and had no unpaid tax assessments from any prior tax period. Our analysis covers the period from 2011 to 2019, a time marked by significant reductions in IRS compliance program resources but preceding the COVID-19 pandemic, which significantly altered taxpayer interactions with the IRS. To assess compliance, we consider only the most recently filed return and current balance due, drawing a 1% random sample of compliant taxpayers each year. This results in a repeated cross-sectional data structure, with approximately 1.2 to 1.3 million observations per year. No taxpayer appears in more than one year, and the total sample size for all years combined is 11.6 million.

Since most compliant taxpayers remain compliant across time (approximately 95%), a substantial portion of the sample is lost when analyzing models that focus on those who fall out of compliance. To address this issue, we generate an alternative 10% sample using a different seed for the random sampling. This approach significantly increases the number of non-compliant taxpayers in the dataset, expanding the sample size from around 600,000 to nearly 6,000,000 observations.

Tax information for the study population during the pilot year was compiled from individual return filings, data on unpaid tax assessments, and information return filings. These datasets provide comprehensive details on income types and amounts, changes in outstanding balances, compliance risk scores, exam classification groups, and other characteristics. We use prior year (t-1) data as controls and predictors in our models for current year (t) outcomes. Our primary dependent variables are current year filing and payment compliance, and for taxpayers who fail to file and pay on time, we examine the magnitude of their outstanding balance due over the course of the current year. Since all sampled taxpayers had a zero-balance due at the start of year t (they were compliant), the balance can only remain at zero or increase if they fail to fully pay during the year.

In addition to analyzing the sample of compliant taxpayers, we construct zip code-level treatment variables to capture broader changes in IRS compliance programs. Our analysis primarily focuses on IRS campus programs that correspond with taxpayers—specifically, delinquent return notices (CP59) aimed at non-filers and ACS communications (LT11, LT16, and LT26 letters) directed to those who have unmet filing or payment obligations. We also include field collection programs, which involve in-person visits to taxpayers with unpaid assessments.

The focus on ACS letters, CP59 notices, and field visits aligns closely with our research objective, as these enforcement actions are directly relevant to taxpayer behavior related to timely filing and full payment. To ensure a comprehensive analysis of IRS enforcement's indirect effects, we also include measures of IRS underreporting compliance: campus and field examinations. Campus exams, conducted remotely, are designed for straightforward issues and simpler cases, whereas field exams are in-person audits for more complex situations, often involving businesses or high-income individuals, with IRS agents reviewing records directly at the taxpayer's location. Field exams, therefore, tend to be more thorough and resource-intensive.

To quantify these enforcements, we aggregate the volume of letters, field visits, and campus and field exams conducted within each zip code. To account for unobserved heterogeneity across zip codes, we incorporate fixed effects using zip code dummies. This approach enhances the accuracy of our results by controlling for time-invariant characteristics that might otherwise confound our estimates.

Table 1 presents the summary statistics for the dependent variable and key treatment variables used in this study, highlighting several important trends. The dependent variable captures taxpayer compliance behavior, categorized into three levels based on filing and payment status:

- **Fully compliant** taxpayers are those who file their tax returns and pay their liabilities by the original due date.

- **Late filers** are taxpayers who miss the filing deadline but still manage to pay the full tax amount by the original return due date.
- **Delinquent** taxpayers include those who fail to pay the full tax amount by the original due date, which includes non-filers and those who still have outstanding liabilities despite filing on time or late.

The summary statistics reveal a noticeable decline in the proportion of compliant taxpayers from 2011 to 2019, accompanied by a corresponding increase in the non-compliant population over the same period. This shift suggests a growing challenge of maintaining compliance levels during the study period.

Table 1: Summary Statistics – Means of Key Variables

	(1) All (2011-2019)	(2) 2011	(3) 2019
Outcome Variables			
Fully Compliant (%)	94.9	95.4	94.1
Late Filer (%)	2.1	1.7	2.6
Delinquent (%)	3.0	2.9	3.3
Δ Balance Due (All Prev. Compliant, $)	157	115	186
Δ Balance Due (Delinquent Only, $)	5213	3873	5666
Treatment Variables			
ACS	172.5	286.3	72.8
CP59	113.0	183.1	87.4
Field collection	11.0	13.6	7.7
Campus exam	9.9	12.4	6.5
Field exam	43.5	56.4	33.4
Control Variables			
Married filing jointly	0.37	0.39	0.36
Log total positive income	10.33	10.24	10.42
Timely filed in past four years	0.73	0.75	0.72
Balance due (before remittance)	0.13	0.11	0.14
% of income under-withheld	-0.07	-0.08	-0.06
≥50% of income not subject to withholding	0.10	0.10	0.10
Observations	10,246,313	1,086,418	1,181,211

Notes: Means of each variable are presented for each category over the entire sample period (2011-2019) in All and separately for the years 2011 and 2019. The treatment variables (ACS, CP59, Field collection, Campus exam, and Field exam) are zip code-level counts. The control variables reflect taxpayer-level measures from the prior tax year. Units are specific to each variable, where applicable.

In addition to changes in taxpayer compliance, our main treatment variables, which represent different enforcement actions—ACS notices, CP59 notices, and field visits—show significant decreases over the study period. Specifically, the average number of ACS notices per zip code fell from 286 in 2011 to 73 in 2019, a reduction of approximately 74%. CP59 notices also declined, dropping from 183 to 87 per zip code, a decrease of about 52%. Field visits saw a similar downward trend, decreasing from 14 to 8 per zip code, representing a 43% reduction. The data also show considerable disparities in the frequency of enforcement actions. On average, 173 ACS notices were sent per zip code annually, which is 53% higher than the average number of CP59 notices. Field visits were even less frequent, with ACS notices being issued 16 times more often than field visits, which averaged only 11 per zip code. These disparities in the frequency and scale of enforcement actions suggest that the marginal impact of each treatment variable on taxpayer compliance may vary substantially.

Overall, the summary statistics underscore the critical role of enforcement actions in shaping tax compliance behavior and highlight the significant variation in the intensity of different enforcement strategies. The sharp decline in enforcement activities over the years raises concerns about the IRS's ability to sustain compliance rates, especially as resource constraints continue to limit its operational capacity.

B. Social Connectedness Index

To accurately assess the indirect effects of IRS enforcement, it is crucial to understand how information about enforcement actions circulates through social networks, which may span geographic areas, preparer networks, or supply chains. For this purpose, we use the Social Connectedness Index (SCI), developed by Bailey et al. (2018). The SCI measures the intensity of connections between zip code pairs using anonymized Facebook friendship data from 2016, a time when approximately two-thirds of all U.S. adults used Facebook (Greenwood et al., 2016), to reflect the density of social connections across the U.S. Given Facebook's extensive user base and a demographic profile that mirrors the general population, the SCI provides a reliable indicator of social networks, offering valuable insights into how social ties influence perceptions of enforcement actions.

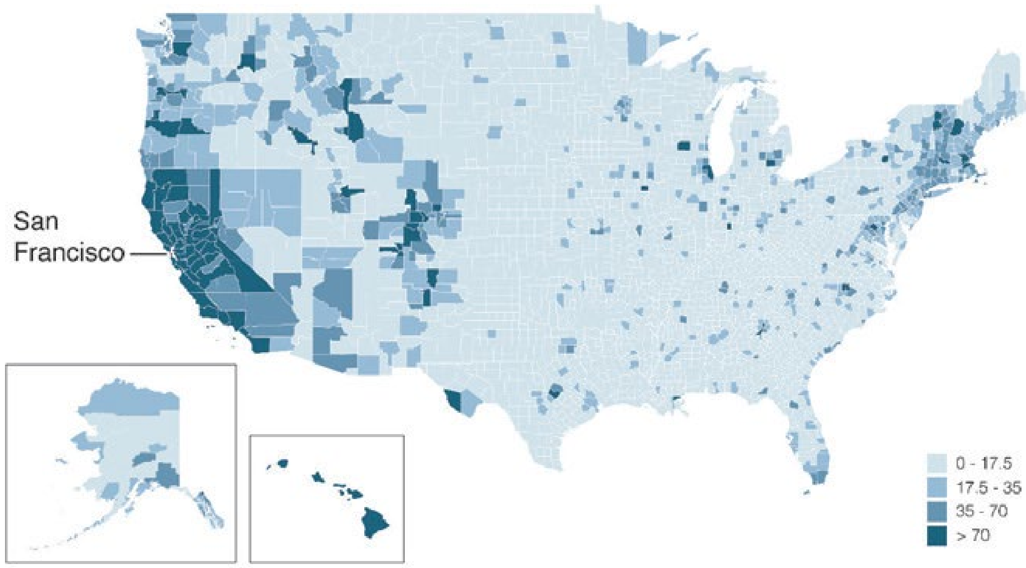
Unlike traditional measures of social proximity that rely on geographic location, the SCI captures actual social connections, offering a more nuanced understanding of how individuals are linked across regions. For example, as depicted in Figure 4, while San Francisco County and Kern County in California have similar population sizes, their social networks are markedly different. San Francisco's connections are dispersed nationally, particularly into the Northeast, while Kern County's network is concentrated on the West Coast, with strong ties to regions such as Oklahoma and Arkansas due to historical migration patterns. Specifically, 57% of Kern County's friendships are within 50 miles, closely matching the U.S. average of 55.4%, whereas only 27% of San Francisco's friendships are within the same range, highlighting its broader social dispersion. By calculating the "relative probability of friendship"—adjusted for the number of Facebook users—the SCI provides a more precise measure of social connectedness that goes beyond simple geographic proximity. This measure is crucial for understanding how perceptions of IRS enforcement actions spread within and across communities, as geographic closeness alone does not fully capture the strength and influence of social ties.

Data Coverage and the Social Connectedness Index

During the study period from 2011 to 2019, an average of 148 million individual tax returns were filed annually across 58,960 zip codes, covering all 50 states and the District of Columbia. This figure includes not only standard geographic zip codes but also PO Box-only zip codes, unique codes for large organizations, and military zip codes. By contrast, according to U.S. Postal Service data from 2024, there are approximately 41,704 standard geographic zip codes in the United States. Our 1% random sampling of individual tax returns reduces the number of zip codes in our dataset to 39,794 out of the 58,960 total zip codes. Crucially, only 0.01% of tax returns are filed in zip codes outside of this sample, ensuring that our data remains highly representative of taxpayer behavior across the U.S.

The SCI, used to capture social connections between zip codes, further limits coverage due to privacy concerns, excluding zip codes with very few users. As a result, the SCI encompasses 22,718 zip codes, representing the zip codes for which our weighted average treatment variables are available. While the exclusion of some zip codes might seem significant, it is important to note that the 22,718 zip codes covered by the SCI account for 97% of all tax returns filed during the 2011 to 2019 period. The remaining 3% of tax returns come from zip codes in remote areas with sparse populations and minimal tax return activity, meaning their exclusion has little impact on the representativeness of our analysis. Therefore, our dataset captures the majority of taxpayer interactions and remains robust for the purposes of our analysis.

(a) Relative Probability of Friendship Link to San Francisco County, CA



(b) Relative Probability of Friendship Link to Kern County, CA

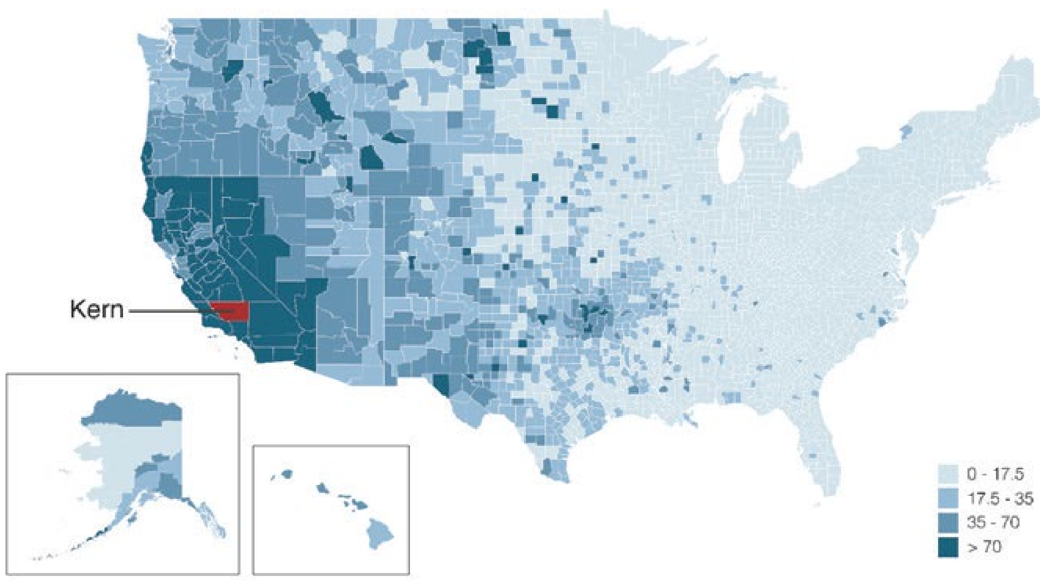


Figure 4. County-Level Friendship Maps

Note: The heat maps illustrate the relative likelihood of a Facebook user in each county j having a friendship connection with San Francisco County, CA (Panel a) and Kern County, CA (Panel b). Darker shades indicate counties where there is a greater likelihood of a friendship connection from a person in the home county (San Francisco or Kern) to county j. The "relative probability of friendship" is derived by dividing the Social Connectedness Index between counties i and j by the product of the total number of Facebook users in both counties. *SOURCE: Bailey et al. (2018).*

Treatment Variables Transformation

To capture the indirect effects of IRS enforcement actions, we transform key treatment variables, ACS notices, CP59 notices, and field visits, using weighted averages based on the SCI. This transformation accounts for how enforcement actions in one zip code may influence taxpayer behavior in socially connected zip codes, reflecting the spread of enforcement perceptions through social networks.

For example, the transformation of ACS notices is calculated as follows:

(1) $$ACS_{jt} = \sum_{k} w_{jk} ACS_{kt}^{raw}$$

where ACS_{jt} represents the weighted average of ACS letters sent to zip code j in year t, w_{jk} is the social connection measure between zip code j and k, and ACS_{kt}^{raw} is the number of ACS letters sent to zip code k in year t. This method enables our models to capture how social connections, rather than geographic proximity alone, shape the dissemination of enforcement perceptions and influence taxpayer behavior. By incorporating SCI-weighted averages, we reflect the indirect effects of enforcement actions as they propagate through connected communities.

Table 2 summarizes the treatment variables after applying the SCI transformation. The overall trends remain similar to the raw data, showing noticeable declines in enforcement actions over time. However, the SCI-weighted variables are larger on average, reflecting the amplifying effect of social connections. Importantly, the standard deviations of the treatment variables decrease significantly after the transformation, indicating that the SCI smooths out extreme variations in the raw data, explained in more detail below. This reduction is because some zip codes that received fewer direct enforcement actions in the raw data are socially connected to others that received more intensive enforcement, allowing for a more accurate measure of the indirect effects through social spillovers.

Table 2: Treatment Variables after Transformation

	All		2011		2019	
Variable	Mean	SD	Mean	SD	Mean	SD
Unweighted ACS	173	280	286	383	73	141
Unweighted CP59	113	182	183	248	87	168
Unweighted Field collection	11	14	14	18	7.7	9.0
Unweighted Campus exam	9.9	15	12	18	6.5	9.4
Unweighted Field exam	44	75	56	97	33	51
SCI-Weighted ACS	191	144.7	319	152	81.6	38.9
SCI-Weighted CP59	124	92.9	206	99.4	100	51.1
SCI-Weighted Field collection	11.9	5.9	15.2	7.1	8.4	3.3
SCI-Weighted Campus exam	49.2	33.4	64.0	42.7	37.2	21.3
SCI-Weighted Field exam	10.8	7.1	13.8	8.3	7.0	3.6

Notes: "Mean" and "SD" denote the mean and standard deviation for the entire sample period (2011-2019) in All and separately for the years 2011 and 2019.

Smoothing Effect of the SCI Transformation

Figure 5 compares the distribution of actual ACS letters sent to various zip codes in the Washington D.C. area with the distribution after the SCI transformation. This comparison highlights two key points. First, the SCI transformation smooths out the varying values of ACS notices across different zip codes. In the left panel, some zip codes received over 1,000 notices, while adjacent zip codes received only a few. This stark variation can be misleading for analyzing indirect effects, as these effects propagate through social connections, which often align with—but do not strictly adhere to—geographic proximity. The right panel, which uses SCI-weighted data, shows a more gradual variation in ACS notices, offering a clearer understanding of how enforcement messages spread through social networks.

Second, the contrasting examples of zip codes 20762 (Joint Base Andrews) and 20742 (University of Maryland) illustrate that geographic proximity alone does not fully explain how enforcement effects propagate. In 2011, zip code 20762 received only about 20 notices, despite surrounding areas receiving over 1,000. Even after the SCI transformation, the low social connectivity of this military base results in a relatively low number of notices. In contrast, zip code 20742, which initially received fewer than 1% of the notices compared to its neighbors, shows almost no difference after the SCI transformation due to its higher social connectivity.

These examples highlight the critical role of social networks—rather than geographic distance alone—in determining how enforcement messages disseminate across regions.

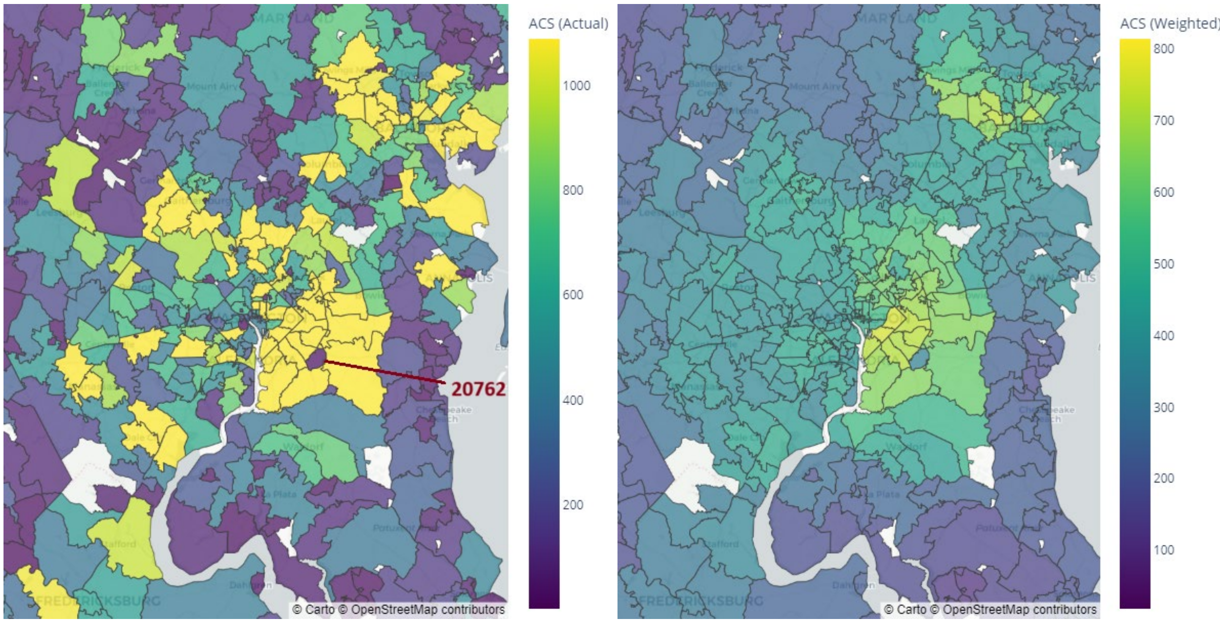


Figure 5. Comparison of Raw and SCI-Weighted ACS Letters in the Washington D.C. Area

Notes: These panels illustrate the distribution of ACS letters across different zip codes in the Washington D.C. area. The left panel displays the raw counts of notices sent in 2011, winsorized at the top 5% level to enhance visual clarity. The right panel presents the ACS notices weighted by the Social Connectedness Index. The color bars indicate the respective ranges for each panel.

C. Regression Modeling Framework

To evaluate the causal impact of IRS enforcement actions on taxpayer behavior, we employ a two-stage econometric approach that accounts for both the extensive and intensive margins of compliance. The first stage estimates the likelihood of behavioral transitions among previously compliant taxpayers, categorizing them into three distinct compliance states: continued full compliance, late filing, and delinquency with an outstanding balance. We use a multinomial logistic model to analyze how exposure to enforcement actions indirectly affects these transitions.

In the second stage, we investigate the financial consequences for taxpayers who enter delinquency, estimating the effect of enforcement actions on the magnitude of outstanding balances. This is accomplished using an ordinary least squares (OLS) regression, where the dependent variable captures changes in the balance due. By integrating these two stages, our framework enables a comprehensive assessment of enforcement effectiveness, distinguishing its role in preventing non-compliance (deterrence effect) and mitigating the financial severity of delinquency (recovery effect).

Two-stage Multinomial Logistic Model for Filing and Payment Compliance

To address potential endogeneity in the ACS, CP59, Field, Campus, and Field Exam variables—where regions with higher non-compliance may experience greater enforcement efforts—we employ a two-stage least squares (2SLS) approach. This method allows us to disentangle the effects of enforcement actions from the reverse causality driven by underlying non-compliance rates. By using the IRS's annual FTE allocations for specific types of enforcement as instrumental variables (IVs), we isolate exogenous variations in enforcement, producing unbiased estimates of enforcement effects on compliance. These FTE allocations are determined administratively and are, therefore, exogenous to taxpayer compliance behaviors, making them ideal instruments for this setting.

First Stage: In the first stage, we model enforcement variables (ACS, CP59, Field, Campus, and Field Exam) for each year t and zip code j, using FTE positions allocated annually to each type of enforcement as instrumental variables (IVs). Unlike ACS and field collection programs, CP59 notices do not have dedicated FTEs. Instead, FTEs allocated to collection enforcement units—such as ACS and field collection—are interchangeably used for CP59 cases as well. To reflect the shared and overlapping nature of IRS collection efforts, we utilize both ACS and field collection FTEs, along with interaction terms, to predict the number of ACS, CP59, and Field interventions.

In contrast, the exam units operate more distinctly from the collection units. Campus and field exams have their own specific FTE allocations, and these are used directly in our models. To capture potential non-linear relationships, such as diminishing or increasing returns from increased staffing, we include quadratic terms for each type of exam-related FTEs. This nuanced modeling approach enables us to better understand how variations in IRS staffing—whether shared among collection units or specific to exams—impact enforcement activities.

The model formulations for the first stage are as follows:

$$ACS_{jt} = \alpha_1 + \beta_1 FTE_t^{ACS} + \beta_2 FTE_t^{Field\ Collection} + \beta_3 FTE_t^{ACS} * FTE_t^{Field\ Collection} + \gamma_{zip} + v_{jt} \tag{2}$$

$$CP59_{jt} = \alpha_2 + \beta_4 FTE_t^{ACS} + \beta_5 FTE_t^{Field\ Collection} + \beta_6 FTE_t^{ACS} * FTE_t^{Field\ Collection} + \gamma_{zip} + v_{jt} \tag{3}$$

$$Field_{jt} = \alpha_3 + \beta_7 FTE_t^{ACS} + \beta_8 FTE_t^{Field\ Collection} + \beta_9 FTE_t^{ACS} * FTE_t^{Field\ Collection} + \gamma_{zip} + v_{jt} \tag{4}$$

$$Campus_{jt} = \alpha_4 + \beta_{10} FTE_t^{Campus\ Exam} + \beta_{11} FTE_t^{Campus\ Exam} * FTE_t^{Campus\ Exam} + \gamma_{zip} + v_{jt} \tag{5}$$

$$FieldExam_{jt} = \alpha_5 + \beta_{12} FTE_t^{Field\ Exam} + \beta_{13} FTE_t^{Field\ Exam} * FTE_t^{Field\ Exam} + \gamma_{zip} + v_{jt} \tag{6}$$

Second Stage: The second stage involves regressing the probability of taxpayer compliance outcomes (P_{ijt}) on the predicted values from the first stage. P_{ijt} is categorized as follows:

- P_{ijt}=0: Fully compliant (filed and paid on time).
- P_{ijt}=1: Filed late but no outstanding balance (paid in full).
- P_{ijt}=2: Has an outstanding balance due at the end of time *t*.

The second stage model is specified as follows:

$$P_{ijt} = F\left(\alpha + \beta_1 \widehat{ACS}_{jt-1} + \beta_2 \widehat{CP59}_{jt-1} + \beta_3 \widehat{Field}_{jt-1} + \beta_4 \widehat{Campus}_{jt-1} + \beta_5 \widehat{FieldExam}_{jt-1} + \sum_k \theta_k X_{ijt-1} + \gamma_{zip} + \eta_{year}\right) + e_{ijt} \tag{7}$$

Here, F() represents the multinomial logit link function. The predicted values from the first stage ($\widehat{ACS}_{jt-1}$, $\widehat{CP59}_{jt-1}$, $\widehat{Field}_{jt-1}$, $\widehat{Campus}_{jt-1}$, and $\widehat{FieldExam}_{jt-1}$) are used as independent variables along with control variables (X_{ijt-1}), zip code (γ_{zip}) and year (η_{year}) fixed effects. This setup leverages administrative FTE allocations as instruments, allowing us to derive causal insights on the effects of enforcement actions while effectively controlling for potential biases from time-invariant regional and temporal factors.

Additional Control Variables: We include a comprehensive set of taxpayer characteristics based on the most recent return filed in previous year (t-1), which are represented in X_{ijt-1}. These control variables are fully listed in Appendix Table A1 and account for differences in compliance behavior, income, and risk characteristics, and include:

- Indicator for married filing jointly status.

- Log transform of total positive income.
- Indicator for filing on-time consecutively for the last four years.
- Indicator for having a balance due (from line 37 of Form 1040, before remittance).
- Under-withholding as a percent of total positive income (balance due/ total positive income), restricted to between -100% and 100%.
- Indicator for 50% or more of income derived from sources that cannot withhold taxes, such as self-employment income.
- Indicators for activity code/audit class indicators and their interactions with the Discriminant Function (DIF) score.

These measures help capture the taxpayer's risk profile, with the DIF score serving as a proxy for reporting compliance and the likelihood of an audit, allowing us to control for potential selection biases. The DIF score is uniquely defined by the activity class of the return, so we also include indicator variables for these classes and interaction terms between them and the DIF score. Because our taxpayers were selected based on being compliant in the prior year, they aren't directly treated by a filing and payment compliance program, and there is no direct measure of the impact of compliance programs in our model. For this population we aim to measure only an indirect impact of these programs.

Linear Model for Change in Balance Due

Our sample of taxpayers begins each year *t* with no outstanding balance due. While most taxpayers will maintain this status throughout the year, some will fall out of compliance and receive a balance due notice. For these individuals, we model the indirect effects of compliance programs on the change in their outstanding balance. This approach allows us to evaluate whether these programs can positively influence compliance by reducing the size of the taxpayer's debt, even if they do not entirely prevent non-compliance. For filers who receive a balance due notice after underpaying, we use the total balance shown on the initial notice. For non-filers, we calculate the balance due based on the information returns provided to the IRS. This framework enables us to estimate the intensive margin for taxpayers who do not remain fully compliant.

We define our outcome variable for the change in balance due, U_{ijt} as follows:

- For taxpayers with P_{ijt}=2, U_{ijt} is the amount of tax not timely filed and paid. For filers, this is the total balance due on the first notice sent to the taxpayer. For non-filers, it is the balance due on a potential substitute for return (SFR) generated through the Case Creation Nonfiler Identification Process
- Otherwise, U_{ijt}=0 (late filers with P_{ijt}=1 or compliant taxpayers with P_{ijt}=0)

Our 1% sample of compliant taxpayers includes about 11.6 million observations, but only around 350,000 (3%) of them ended up with an outstanding tax debt ($U_{ijt} > 0$). Because this number is insufficient for robust analysis across the comprehensive set of zip codes and multiple years used in our two-stage approach, we employ an alternative 10% sample for the balance due model. This adjustment increases the sample size to approximately 3.5 million previously compliant taxpayers who later accrued outstanding tax debts.

We use this 10% sample to run a linear model for the change in balance due, focusing on the 3.5 million taxpayers with an outstanding balance. The dependent variable, U_{ijt}, is log-transformed to mitigate bias caused by skewed unpaid tax amounts with extreme outliers. We calculate U_{ijt} using tax year *t*-1, which is filed in year t. Following the two-stage approach outlined in models (2)-(7), we run the following ordinary least squares (OLS) regression for taxpayers with P_{ijt}=2:

$$\log(U_{ijt}) = \alpha + \beta_1 \widehat{ACS}_{jt-1} + \beta_2 \widehat{CP59}_{jt-1} + \beta_3 \widehat{Field}_{jt-1} + \beta_4 \widehat{Campus}_{jt-1} + \beta_5 \widehat{FieldExam}_{jt-1} + \sum_k \theta_k X_{ijt-1} + \gamma_{zip} + \eta_{year} + e_{ijt} \tag{8}$$

Similar to model (7), model (8) predicts the unpaid assessment amount U_{ijt} for non-compliant taxpayer *i* in zip code *j* in year *t*. It is regressed on the predicted values of endogenous variables, $\widehat{ACS}_{jt-1}$, $\widehat{CP59}_{jt-1}$,$\widehat{Field}_{jt-1}$, $\widehat{Campus}_{jt-1}$, and $\widehat{FieldExam}_{jt-1}$ with the same control variables X_{ijt-1} as in model (5), along with zip code (γ_{zip}) and year (η_{year}) fixed effects to account for omitted variables that may influence U_{ijt}. Using the two-stage process in model (8) also has the same advantages as with model (7), better controlling for unobserved factors through the incorporation of fixed effects and addressing potential endogeneity between the zip code-level indirect treatments and U_{ijt} using the IRS enforcement budget as an instrumental variable.

Further tests confirm that the SCI-based model outperforms alternative models that rely on simple geographical distances or unweighted counts of enforcement actions to replace the

SCI-based treatment variables. Comprehensive results and comparisons from these additional model tests, presented in the Appendix Table A2, substantiate the effectiveness of using the SCI to capture the indirect effects of IRS enforcement strategies.

IV. Results

A. Two-Stage Multinomial Logistic Model for Filing and Payment Compliance

Model Results and Interpretation

Our two-stage model utilizes FTE allocations as instrumental variables in the first stage to predict enforcement variables, followed by a second stage that models compliance outcomes based on these predicted values. Table 3 presents the first stage regression results, which show that FTE allocations positively affect the number of enforcements, with a strong model fit indicated by the R-squared and F-statistics. The negative interaction term between collection FTEs reflects their interchangeable allocation, while the quadratic terms for exam FTEs suggest diminishing returns, consistent with typical labor input-output relationships.

Table 4 presents the results from the multinomial compliance model. The findings align with intuitive expectations—positive coefficients in the compliant category (P_{ijt}=0) suggest that increased enforcement efforts improve compliance, while negative coefficients for the non-compliant categories (P_{ijt}=1 and P_{ijt}=2) indicate a reduction in the likelihood of non-compliance. Among the three enforcement programs, ACS letters demonstrate the strongest influence, followed by CP59 notices. Although the sample consists of generally compliant taxpayers, the model reveals that increased enforcement—especially through ACS letters—has a significant preventative effect, enhancing voluntary compliance rates. CP59 notices similarly contribute to compliance improvements, though to a lesser extent than ACS letters. Field collection interventions, while impactful, exhibit a more modest effect in comparison to the other two programs.

Table 3. First-Stage Regression Results for Two-State Least Squares Model

Variable	ACS	CP 59	Field Collection	Campus Exam	Field Exam
Intercept	-312.2*** (8.14)	-213.4*** (7.33)	2.88*** (0.38)	-47.49*** (3.40)	-0.99 (0.94)
ACS FTE	0.129*** (0.002)	0.081*** (0.001)	0.0007*** (0.00006)	-	-
Field Collection FTE	0.076*** (0.001)	0.029*** (0.000)	0.0015*** (0.00003)	-	-
ACS FTE × Field Collection FTE	-0.019*** (0.0004)	-0.005*** (0.0002)	-0.0003*** (0.00001)	-	-
Campus Exam FTE	-	-	-	0.043*** (0.000)	-
$Campus\ Exam\ FTE^2$	-	-	-	-0.053*** (0.0003)	-
Field Exam FTE	-	-	-	-	0.001*** (0.00004)
$Field\ Exam\ FTE^2$	-	-		-	-0.0001*** (0.00001)
Zip Code FE	Y	Y	Y	Y	Y
R-squared	0.821	0.816	0.911	0.900	0.912
F-statistic	211.0	49.67	152.1	124.0	76.26
Number of Observations			185,593		

Note: Standard errors in parentheses. *, **, *** indicate significance levels of $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively. For readability, interaction terms (e.g., ACS FTE x Field Collection FTE) are multiplied by 1,000 and quadratic terms are multiplied by 10,000. All other coefficients are reported in their original scale.

Table 4. Selected Parameter Estimates for Two-Stage Multinomial Compliance Model
(P_{ijt}=0: compliant, 1: non-compliant no balance due, 2: non-compliant with balance due)

Variable	P_{ijt}=0	P_{ijt}=1	P_{ijt}=2
Intercept	0.211 *** (0.016)	-0.074 *** (0.026)	-0.137 *** (0.022)
ACS weighted average	5.539 *** (0.005)	-2.108 *** (0.009)	-3.431 *** (0.008)
CP59 weighted average	3.155 *** (0.003)	-1.202 *** (0.005)	-1.954 *** (0.004)
Field collection weighted average	0.278 *** (0.000)	-0.102 *** (0.000)	-0.176 *** (0.000)
Campus exam weighted average	1.036 *** (0.001)	-0.371 *** (0.001)	-0.665 *** (0.001)
Field exam weighted average	0.243 *** (0.000)	-0.089 *** (0.000)	-0.154 *** (0.000)
Married filing jointly	0.186 *** (0.002)	-0.179 *** (0.004)	-0.007 ** (0.003)
Log total positive income	-0.128 *** (0.001)	0.011 *** (0.002)	0.117 *** (0.001)
Timely filed in past four years	0.591 *** (0.002)	-0.322 *** (0.003)	-0.269 *** (0.003)
Balance due (before remittance)	-0.136 *** (0.003)	-0.049 *** (0.005)	0.185 *** (0.004)
Percent of income under-withheld	-1.478 *** (0.010)	-0.142 *** (0.016)	1.620 *** (0.015)
50% or more of income not subject to withholding	-0.101 *** (0.003)	-0.008 (0.005)	0.110 *** (0.005)
Year Fixed Effects	Y	Y	Y
Zip Code Fixed Effects	Y	Y	Y
Number of Observations		10,246,313	

Note: Standard errors in parentheses. *, **, *** indicate significance levels of $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively.

Impact Analysis

The ACS intervention demonstrates the most substantial influence on compliance among the programs studied. Over the study period from 2011 to 2019, ACS notices were sent to approximately 45,000 zip codes. In comparison, CP59 notices and field collections were administered to around 40,000 and 33,000 zip codes, respectively. Additionally, the frequency of ACS treatments per zip code significantly outpaces that of CP59 and field visits. On average, 173 ACS letters were sent per zip code annually, compared to 113 CP59 notices and just 11 field visits per zip code each year.

The disparity in both the breadth and intensity of enforcement efforts leads to differing impacts across programs. Our findings emphasize that the wide reach and frequent interactions of the ACS program are particularly effective in enhancing voluntary compliance. These indirect effects, which spread through social networks, extend the impact of enforcement actions beyond directly treated individuals. By ensuring compliance programs have sufficient resources to contact taxpayers, the IRS can amplify the spread of compliant behavior across a wider population. In contrast, direct effects are limited to those directly treated and follow a different dynamic. For instance, while field visits are more limited in scope, they may exert a stronger direct effect due to the intensity of in-person contact, prompting immediate compliance.

Average Marginal Effects and Impact of Increased Enforcement Levels

The average marginal effects from the multinomial model, shown in Table 5, convert log odds into probabilities, offering a clearer interpretation of the enforcement programs' impact on compliance. An increase of 1,000 ACS letters leads to significant reductions in both late filings and delinquencies, indicating substantial improvements in compliance. Similarly, increases in CP59 notices and field visits also lower non-compliance rates, though to a lesser degree. The results highlight the varying effectiveness of these enforcement tools, with ACS letters proving to be particularly powerful in fostering taxpayer compliance.

Table 6 expands on these findings by showing the marginal effects of a 10% increase in each program's enforcement levels. A 10% increase in ACS letters is associated with a 0.3 percentage point decrease in late filings and a 0.5 percentage point decrease in delinquencies, corresponding to 15% and 17% reductions, respectively. CP59 notices also yield positive effects, with a 10% increase reducing late filings by 0.1 percentage points (5% improvement) and delinquencies by 0.2 percentage points (6% decrease). Field collection visits have a more modest effect, highlighting that while effective, their reach is more limited compared to the broader, more frequent ACS letters and CP59 notices.

Table 5. Average Marginal Effects for Two-Stage Multinomial Compliance Model†
(P_{ijt}=0: compliant, 1: non-compliant no balance due, 2: non-compliant with balance due)

Variable	P_{ijt}=0	P_{ijt}=1	P_{ijt}=2
ACS weighted average	0.397 ***	-0.149 ***	-0.249 ***
	(0.003)	(0.001)	(0.002)
CP59 weighted average	0.226 ***	-0.085 ***	-0.142 ***
	(0.001)	(0.000)	(0.001)
Field collection weighted average	0.020 ***	-0.007 ***	-0.013 ***
	(0.000)	(0.000)	(0.000)
Campus exam weighted average	0.074 ***	-0.027 ***	-0.047 ***
	(0.000)	(0.000)	(0.000)
Field exam weighted average	0.017 ***	-0.006 ***	-0.011 ***
	(0.000)	(0.000)	(0.000)
Married filing jointly	0.012 ***	-0.007 ***	-0.005 **
	(0.000)	(0.000)	(0.000)
Log total positive income	-0.009 ***	0.003 ***	0.007 ***
	(0.000)	(0.000)	(0.000)
Timely filed in past four years	0.049 ***	-0.021 ***	-0.028 ***
	(0.000)	(0.000)	(0.000)
Balance due (before remittance)	-0.011 ***	0.002 ***	0.010 ***
	(0.000)	(0.000)	(0.000)
Percent of income under-withheld	-0.112 ***	0.025 ***	0.087 ***
	(0.001)	(0.000)	(0.002)
50% or more of income not subject to withholding	-0.008 ***	0.002 ***	0.006 ***
	(0.000)	(0.000)	(0.000)
Year Fixed Effects	Y	Y	Y
Zip Code Fixed Effects	Y	Y	Y
Number of Observations		10,246,313	

Note: Standard errors in parentheses. *, **, *** indicate significance levels of $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively.

Table 6. Marginal Effect Estimates for 10% Increase in Program Levels

	Δ Probability for Late Filers	Δ Probability for Delinquent Cases
ACS Letters	-0.3	-0.5
CP59 Notices	-0.1	-0.2
Field Collection	-0.001	-0.002
Campus Exam	-0.02	-0.03
Field Exam	-0.0008	-0.001

Our results confirm the differential effectiveness of IRS compliance programs. ACS letters, due to their broad distribution and frequency, are especially potent in encouraging taxpayer compliance. In contrast, CP59 notices and field visits, while effective, have a more limited reach. These findings suggest that strategic resource allocation focusing on extensive and frequent outreach, particularly through ACS, is critical for enhancing voluntary compliance. Policymakers can use these insights to optimize enforcement efforts and refine program designs for greater efficiency.

B. Linear Model for Change in Balance Due

OLS Results

Table 7 shows the OLS results for the change in outstanding balance due, shown in Equation (8). Table 7 reveals patterns consistent with our findings from the filing and payment compliance models. Specifically, the parameter estimates for ACS letters and CP59 notices are negative and statistically significant, indicating that enforcement actions contribute to reducing unpaid tax balances, even when they do not prevent taxpayers from becoming delinquent altogether.

The results indicate that ACS letters exert the greatest impact in reducing the outstanding balance due, followed by CP59 notices. Field collection interventions, while statistically significant, show only a marginal effect, with significance at the 10% level. This suggests that although field visits are a more intensive enforcement action and may generate substantial direct effects, their overall indirect impact on reducing balances is minimal compared to the broader influence of ACS letters and CP59 notices. These findings underscore the effectiveness of widespread, less resource-intensive interventions in mitigating delinquent balances through indirect channels.

Table 7. Selected Parameter Estimates for Linear Model of Change in Balance Due

Variable	$Log(U_{ijt})$
Intercept	5.525*** (0.017)
ACS weighted average	-0.029*** (0.006)
CP59 weighted average	-0.015*** (0.003)
Field collection weighted average	-0.001** (0.000)
Campus exam weighted average	-0.000 (0.001)
Field exam weighted average	-0.001*** (0.000)
Married filing jointly	0.068*** (0.002)
Log total positive income	0.202*** (0.001)
Timely filed in past four years	-0.177*** (0.002)
Balance due (before remittance)	-0.184*** (0.003)
Percent of income under-withheld	0.334*** (0.011)
50% or more of income not subject to withholding	0.031*** (0.003)
Year Fixed Effects	Y
Zip Code Fixed Effects	Y
Number of Observation	3,286,146

Note: Standard errors in parentheses. *, **, *** indicate significance levels of $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively.

C. Nationwide Delinquent Balances and Enforcement Impacts

Nationwide Estimates of Delinquent Balances

To estimate the nationwide balance due for previously compliant taxpayers who became delinquent, we employ two complementary approaches, both of which yield consistent estimates of approximately \$19.6 billion in yearly delinquent balances for the period 2011-2019. The first approach uses a 10% sample of taxpayers who were compliant at the start of the year but ended the year with a delinquent balance. The aggregated balances from this sample is scaled up to represent the entire population:

$$\text{(9)} \qquad One\ Year\ Delinquent\ Balance = \left(\sum_i balance_i\right) * \frac{10}{9}$$

The second approach leverages a 1% sample of taxpayers to estimate the national delinquent balance by combining three components: the average number of compliant taxpayers (TCP ≈ 129 million), the probability of transitioning to delinquency ($P_{ijt} = 2$, ≈ 3%), and the expected balance among delinquents ($E\left[Balance \middle| P_{ijt} = 2\right]$ ≈ \$5,000):

$$\text{(10)} \qquad One\ Year\ Delinquent\ Balance = TCP * P_2 * E\left[Balance \middle| P_{ijt} = 2\right]$$

Both approaches provide independent but consistent estimates of the annual nationwide delinquent balance for taxpayers who were compliant at the start of the year.

Total National Impact of Enforcement Actions

To quantify the effect of enforcement actions, we estimate the Total National Impact (TNI) of interventions, including ACS letters, CP59 notices, and field collections. The TNI is calculated by combining the changes in extensive and intensive margins:

$$\text{(11)} \qquad TNI = TCP * (\, \Delta P_2 * E[\,Balance \mid P_{ijt} = 2\,] + P_2 * \Delta E[\,Balance \mid P_{ijt} = 2\,]),$$

where:

- ΔP_2 represents the change in the probability of delinquency, derived from the multinomial logit regression.
- $\Delta E[\,Balance \mid P_{ijt} = 2\,]$ represents the change in the expected balance among delinquents, derived from our OLS regression.

The change in the intensive margin, $\Delta E[\,Balance \mid P_{ijt} = 2\,])$, is calculated using the coefficient β from the OLS regression of $log\ (U_{ijt})$ on enforcement actions:

$$\text{(12)} \qquad \Delta E[\,Balance \mid P_{ijt} = 2\,]) = E[\,Balance \mid P_{ijt} = 2\,] * (e^{\beta} - 1)$$

Applying this framework, a 10% increase in ACS interventions leads to an estimated \$3.2 billion reduction in newly created delinquent balance, representing a 16% decrease. This estimate applies specifically to taxpayers who were fully compliant in the prior year but became delinquent in the current year. The impact reflects both a reduced probability of transitioning into delinquency and a reduction in the amount of unpaid balances accrued by those who do become delinquent. A similar 10% increase in CP59 notices results in a $1.3 billion decrease (7%), while a 10% increase in field visits yields a much smaller reduction of $11 million (0.06%).

These results demonstrate the efficacy of broad and frequent interventions, such as ACS and CP59 notices, in reducing outstanding balances. In contrast, field collections—despite their direct and intensive nature—have limited indirect impact on individual taxpayer balances. It is noteworthy that field collections are likely more impactful for business taxpayers, consistent with Boning et al. (2019).

Confidence Intervals and Delta Method

We compute the confidence intervals (CIs) for TNI using the delta method, which provides a first-order approximation of variance for non-linear functions of estimated parameters. Specifically, the variance of TNI is expressed as:

$$Var(TNI) = \left(TCP * E\left[Balance \mid P_{ijt} = 2\right] * SE_{\Delta P_2}\right)^2 + \left(TCP * P_2 * SE_{\Delta E[Balance|P_{ijt}=2]}\right)^2,$$

where:

- $SE_{\Delta P_2}$ is the standard error of ΔP_2,
- $SE_{\Delta E[Balance|P_{ijt}=2]}$ is the standard error of $\Delta E\left[Balance \mid P_{ijt} = 2\right]$)

Using the variance, the 95% confidence interval for TNI is calculated as:

$$CI_{TNI} = TNI \pm 1.96\sqrt{Var(TNI)}$$

Table 8. Reductions in Delinquent Balances from a 10% Increase in Enforcement

Compliance Program	Dollar reduction ($B)	Percentage Reduction (%)
ACS Letters	-3.18 [-3.24, -3.12]	-16.2 [-16.5, -15.9]
CP59 Notices	-1.34 [-1.36, -1.33]	-6.85 [-6.93, -6.77]
Field Collection	-0.01 [-0.01, -0.01]	-0.06 [-0.06, -0.06]
Campus Exam	-0.172 [-0.173, -0.172]	-0.88 [-0.88, -0.88]
Field Exam	-0.009 [-0.009, -0.009]	-0.04 [-0.04, -0.04]

V. Conclusion

This study offers a rigorous approach to estimate voluntary compliance effects of IRS enforcement strategies, focusing on ACS letters, CP59 notices, and field collection interventions, and their influence on filing and payment compliance for individual taxpayers. By employing a two-stage multinomial logistic model in combination with the SCI, our analysis underscores the significant role these programs play in maintaining and enhancing voluntary compliance, particularly through their indirect effects across various taxpayer segments.

Our results indicate that ACS letters have the most pronounced impact on promoting voluntary compliance filing and payment obligations. This is reflected by their extensive coverage, reaching approximately 45,000 zip codes and averaging 173 letters per zip code annually. In contrast, CP59 notices and field collections, though impactful, show less influence, indicating their relatively narrower reach and lower frequency of interaction. Our findings emphasize the importance of strategic outreach, where programs with broad reach and consistent interaction are notably effective in fostering voluntary compliance through indirect channels.

Furthermore, the analysis of average marginal effects emphasizes the substantial benefits of even modest increases in enforcement. A 10% increase in ACS letters is associated with significant reductions in both late filings and delinquencies, demonstrating the effectiveness of widespread, targeted enforcement actions. This suggests that a well-distributed approach can yield meaningful improvements in taxpayer behavior, enhancing overall compliance rates.

A key insight from this study is the heterogeneity in enforcement effectiveness across different regions, shaped in part by social dynamics. Our findings suggest that compliance responses to enforcement actions tend to be stronger in areas with higher levels of social connectedness, implying that community networks may facilitate the transmission of compliance-related information and behavioral norms. Understanding these dynamics can help inform broader discussions on optimizing enforcement strategies without altering fundamental allocation principles.

From an economic perspective, the fiscal impact of enhanced enforcement is substantial. Our analysis, utilizing a combination of multinomial logit and linear regression models, reveals that a 10% increase in ACS interventions is linked to a $3.2 billion reduction in newly accrued delinquent balances among previously compliant taxpayers—an approximately 16% decrease. Similarly, a 10% increase in CP59 notices yields a $1.3 billion reduction (7%), while a comparable increase in field visits produces a $11 million reduction (0.06%). These figures highlight the significant fiscal returns that can be realized through strategic improvements in resources for IRS enforcement activities. These indirect effects are in additional to the substantial direct treatment effects of filing and payment compliance programs.

This study underscores the critical role of indirect effects in IRS enforcement strategies and provides actionable insights for policymakers to refine program designs. We plan to extend our approach to estimate indirect effects for business taxpayer. Future research and policy efforts should continue to explore these dynamics to deepen our understanding of enforcement spillover effects and inform the development of evidence-based compliance strategies.

References

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* vol. 32, no. 3, pp. 259-280 (2018)

Bloomquist, Kim. "Incorporating Indirect Effects in Audit Case Selection: An Agent-Based Approach." *Internal Revenue Service* (2012)

Boning, William C., John Guyton, Ronald Hodge, and Joel Slemrod. "Heard it through the grapevine: The direct and network effects of a tax enforcement field experiment on firms." *Journal of Public Economics* vol. 190, issue C (2019)

Datta, Saurabh, Stacy Orlett, and Alex Turk. "Individual Nonfilers and IRS-Generated Tax Assessments: Revenue and Compliance Impacts of IRS Substitute Assessments When Taxpayers Don't File." *Internal Revenue Service* (2015)

Greenwood, Shannon, Andrew Perrin, and Maeve Duggan. "Social Media Update 2016." Retrieved from https://www.pewresearch.org/internet/2016/11/11/social-media-update-2016/. *Pew Research Center*, (2016)

López-Luzuriaga, Andrea and Carlos G. Scartascini. "Willing but Unable to Pay?: The Role of Gender in Tax Compliance." IDB Publications (Working Paper) 12983, *Inter-American Development Bank* (2023)

Internal Revenue Service. "Data Book, 2010." *Publication 55-B*, Washington, DC (March 2011)

Internal Revenue Service. "Data Book, 2019." *Publication 55-B*, Washington, DC (June 2020)

Internal Revenue Service. "Data Book, 2023." *Publication 55-B*, Washington, DC (April 2024)

Internal Revenue Service. "Tax Gap Projections for Tax Year 2022." *Publication 5869 (Rev. 10-2024)*, Washington, DC (October 2024)

Internal Revenue Service. "Understanding Your CP59 Notice." Retrieved from https://www.irs.gov/individuals/understanding-your-cp59-notice. Accessed January 2025.

Internal Revenue Service. "Understanding Your LT11 Notice." Retrieved from https://www.irs.gov/individuals/understanding-your-lt11-notice-or-letter-1058. Accessed January 2025.

Internal Revenue Service. "Understanding Your LT16 Notice." Retrieved from https://www.irs.gov/individuals/understanding-your-lt16-notice. Accessed January 2025.

Internal Revenue Service. "Understanding Your LT26 Notice." Retrieved from https://www.irs.gov/individuals/understanding-your-lt26-notice. Accessed January 2025.

Internal Revenue Service. "IRS Ends Unannounced Revenue Officer Visits to Taxpayers." Retrieved from https://www.irs.gov/newsroom/irs-ends-unannounced-revenue-officer-visits-to-taxpayers. Accessed January 2025.

Treasury Inspector General for Tax Administration (TIGTA). "Trends in Compliance Activities Through Fiscal Year 2022." *Report No*. 2024-300-011 (2023)

Turk, Alex, and Terry Ashley. "Accounts Receivable Resolution and the Impact of Lien Filing Policy on Sole Proprietor Businesses." *2002 Federal Forecasters Conference Proceedings*, pp. 323-332 (2002)

Turk, Alex, John Iuranich, Stacy Orlett, and Saurabh Datta. "Resolving Unpaid Taxes and the Notice of Federal Tax Lien: Evidence from the Fresh Start Initiative." *Internal Revenue Service* (2016)

Appendix

This appendix presents additional information on the datasets constructed for the analysis and full regression results.

Table A1. Variable Descriptions

Name	Description
Time Trend	Linear trend line, increases by one each year
CP59 coverage rate	Total number of taxpayers receiving CP59 notices for year *t*-1 divided by the total number of taxpayers with delinquent accounts (balance due in collections data) in year *t*-1
ACS letter coverage rate	Total number of taxpayers receiving ACS letters LT11, LT16, or LT26 for year *t*-1 divided by the total number of taxpayers with delinquent accounts (balance due in collections data) in year *t*-1
Field coverage rate	Total number of taxpayers in field collection status at any point in year *t*-1 divided by the total number of taxpayers with delinquent accounts (balance due in collections data) in year *t*-1
Married filing jointly	Indicator for married filing jointly filing status on most recent return, filed in year *t*-1
Log total positive income	Natural log transformation of total positive income (amount of income excluding losses) from most recent return, filed in year *t*-1
Timely filed in past four years	Indicator for taxpayers who fully paid and filed timely in the four most recent years, including years *t*-1, *t*-2, *t*-3, and *t*-4
Balance due (before remittance)	Indicator for taxpayers who had an amount greater than or equal to $100 on the "Amount you owe" line from the most recent return, filed in year *t*-1
% of income under-withheld	Ratio of balance due amount ("Amount you owe" line) to total positive income from most recent return, filed in year *t*-1, capped at -1 (cases with refunds equal or greater than total positive income) and 1 (cases with balance due on filing greater than or equal to total positive income)
50% or more of income not subject to withholding	Indicator for taxpayers with a ratio of income not subject to withholding (e.g., farm income from Schedule F, business income from Schedule C, etc.) to total income greater than 0.5 for the most recent return, filed in year *t*-1
Activity code 266	Indicator for taxpayers in activity code (examination class) 266 (Forms 1040PR/1040SS) on the most recent return, filed in *t*-1
Activity code 270	Indicator for taxpayers in activity code (examination class) 270 (returns with earned income tax credit, total positive income below $200,000 and Schedule C/F gross receipts below $25,000 or not present) on the most recent return, filed in *t*-1
Activity code 271	Indicator for taxpayers in activity code (examination class) 271 (returns with earned income tax credit, total positive income below $200,000 and Schedule C/F gross receipts $25,000 or more) on the most recent return, filed in *t*-1
Activity code 272	Indicator for taxpayers in activity code (examination class) 272 (returns with no earned income credit, total positive income below $200,000 and no Schedule C/E/F or Form 2106) on the most recent return, filed in *t*-1.
Activity code 273	Indicator for taxpayers in activity code (examination class) 273 (returns with no earned income credit, total positive income below $200,000 and with Schedule E or Form 2106 but no Schedule C/F) on the most recent return, filed in *t*-1
Activity code 274	Indicator for taxpayers in activity code (examination class) 274 (returns with no earned income credit, total positive income below $200,000 and non-farm business with Schedule C/F receipts below $25,000) on the most recent return, filed in *t*-1

Activity code 275	Indicator for taxpayers in activity code (examination class) 275 (returns with no earned income credit, total positive income below $200,000 and non-farm business with Schedule C/F receipts $25,000-$99,999) on the most recent return, filed in *t*-1
Activity code 276	Indicator for taxpayers in activity code (examination class) 276 (returns with no earned income credit, total positive income below $200,000 and non-farm business with Schedule C/F receipts $100,000-$199,999) on the most recent return, filed in *t*-1
Activity code 277	Indicator for taxpayers in activity code (examination class) 277 (returns with no earned income credit, total positive income below $200,000 and non-farm business with Schedule C/F receipts $200,000 or more) on the most recent return, filed in *t*-1
Activity code 278	Indicator for taxpayers in activity code (examination class) 278 (returns with no earned income credit, total positive income below $200,000 and farm business not classified elsewhere) on the most recent return, filed in *t*-1
Activity code 279	Indicator for taxpayers in activity code (examination class) 279 (returns with no earned income credit, with Schedule C/F and total positive income $200,000-$999,999) on the most recent return, filed in *t*-1
Activity code 280	Indicator for taxpayers in activity code (examination class) 280 (returns with no earned income credit, no Schedule C/F and total positive income $200,000-$999,999) on the most recent return, filed in *t*-1
Activity code 281	Indicator for taxpayers in activity code (examination class) 281 (returns with no earned income credit and total positive income $1,00,000 or more) on the most recent return, filed in *t*-1. Note that activity code 281 is dropped from the models and serves as the reference category for the series of activity code indicator variables
Activity code*DIF	Interaction term for each activity code indicator and the Discriminant Index Function (DIF) score, which ranks the likelihood of tax changes for taxpayers in the event of an audit and is modeled separately for each activity code. The DIF score can take on positive and negative values, and may be thought of as a risk indicator, but only has meaning in context
Year X	Dummy variable for year X
CP59 weighted average	Used in two-stage models as an alternative for CP59 coverage rate, number of CP59 notices in a specific zip code, weighted by SCI index, distance, or unweighted, as described in equation #. For the unweighted models, a log transformation is applied to address skewness
ACS weighted average	Used in two-stage models as an alternative for ACS letter coverage rate, number of ACS letters in a specific zip code, weighted by SCI index, distance, or unweighted, as described in equation #. For the unweighted models, a log transformation is applied to address skewness
Field collection weighted average	Used in two-stage models as an alternative for field coverage rate, number of taxpayers in field collection in a specific zip code, weighted by SCI index, distance, or unweighted, as described in equation #. For the unweighted models, a log transformation is applied to address skewness
Zip Code X	Dummy variable for zip code X (parameter estimates not shown, as zip codes number in the tens of thousands)

Table A2. Full Parameter Estimates for Two-Stage Logistic Compliance Model†
Response Variable: P_{ijt} (0: compliant, 1: non-compliant)

Variable	SCI Weighted	Distance Weighted	Unweighted
	N=11,616,809	N=11,616,809	N=11,616,809
Intercept	-4.762*** (0.027)	-5.236*** (0.029)	-4.744*** (0.008)
ACS weighted average	-1.367*** (0.009)	-0.081*** (0.007)	-0.037*** (0.002)
CP59 weighted average	-0.753*** (0.005)	-0.039*** (0.004)	-0.033*** (0.002)
Field collection weighted average	-0.066*** (0.000)	-0.002*** (0.000)	-0.015*** (0.002)
Married filing jointly	-0.260*** (0.003)	-0.243*** (0.004)	-0.237*** (0.003)
Log total positive income	0.236*** (0.001)	0.235*** (0.002)	0.233*** (0.002)
Timely filed in past four years	-0.884*** (0.003)	-0.873*** (0.003)	-0.876*** (0.003)
Balance due (before remittance)	0.232*** (0.004)	0.232*** (0.004)	0.235*** (0.004)
% of income under-withheld	2.576*** (0.016)	2.603*** (0.016)	2.582*** (0.016)
50% or more of income not subject to withholding	0.210*** (0.005)	0.202*** (0.005)	0.208*** (0.005)
Activity code 266	0.291*** (0.094)	0.588 (0.425)	0.105 (0.088)
Activity code 270	0.830*** (0.025)	0.520*** (0.024)	0.509*** (0.023)
Activity code 271	1.029*** (0.039)	0.840*** (0.039)	0.829*** (0.038)
Activity code 272	0.543*** (0.024)	0.249*** (0.023)	0.242*** (0.022)
Activity code 273	0.632*** (0.025)	0.339*** (0.023)	0.340*** (0.022)
Activity code 274	1.002*** (0.024)	0.700*** (0.023)	0.694*** (0.022)
Activity code 275	1.068*** (0.027)	0.747*** (0.027)	0.739*** (0.026)
Activity code 276	1.106*** (0.055)	0.458*** (0.059)	0.437*** (0.056)
Activity code 277	1.419*** (0.055)	0.804*** (0.059)	0.747*** (0.057)

Activity code 278	0.655*** (0.029)	0.358*** (0.029)	0.358*** (0.029)
Activity code 279	0.416*** (0.026)	0.083*** (0.025)	0.079*** (0.024)
Activity code 280	0.779*** (0.026)	0.457*** (0.025)	0.446*** (0.024)
Activity code 266*DIF	0.005*** (0.000)	-0.001 (0.002)	0.003*** (0.000)
Activity code 270*DIF	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 271*DIF	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Activity code 272*DIF	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 273*DIF	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 274*DIF	0.002*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 275*DIF	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 276*DIF	0.001*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Activity code 277*DIF	0.000*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 278*DIF	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)
Activity code 279*DIF	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Activity code 280*DIF	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Year 2012	-0.134*** (0.006)	-0.069*** (0.007)	-0.087*** (0.006)
Year 2013	-0.242*** (0.006)	-0.075*** (0.006)	-0.112*** (0.006)
Year 2014	-0.371*** (0.006)	-0.051*** (0.006)	-0.120*** (0.006)
Year 2015	-0.342*** (0.006)	0.012** (0.006)	-0.064*** (0.007)
Year 2016	-0.278*** (0.005)	0.121*** (0.006)	0.036*** (0.007)
Year 2017	-0.265*** (0.005)	0.188*** (0.006)	0.089*** (0.007)
Year 2018	-0.339*** (0.005)	0.142*** (0.006)	0.043*** (0.007)
Year 2019	-0.383*** (0.005)	0.146*** (0.006)	0.041*** (0.007)

† Standard errors in parentheses.
Note: *** indicates significance at the 0.01 level in a two-tailed Z-test
** indicates significance at the 0.05 level in a two-tailed Z-test
* indicates significance at the 0.10 level in a two-tailed Z-test
This model simplifies the multinomial framework into a logistic model with P_{ijt} only taking values of 0 for compliance and 1 for non-compliance, to better highlight the comparison between alternative approaches to weighting connections between zip codes.

Table A3. Full Parameter Estimates for Two-Stage Multinomial Compliance Model†
Response Variable: P_{ijt}

(0: compliant, 1: non-compliant no balance due, 2: non-compliant with balance due)

Variable	P_{ijt}=0 N=11,616,809		P_{ijt}=1 N=11,616,809		P_{ijt}=2 N=11,616,809	
Intercept	0.208 ***	(0.016)	-0.072 ***	(0.026)	-0.135 ***	(0.022)
ACS weighted average	5.547 ***	(0.006)	-2.095 ***	(0.009)	-3.452 ***	(0.008)
CP59 weighted average	3.078 ***	(0.003)	-1.152 ***	(0.005)	-1.926 ***	(0.004)
Field collection weighted average	0.275 ***	(0.000)	-0.100 ***	(0.000)	-0.175 ***	(0.000)
Married filing jointly	0.187 ***	(0.002)	-0.179 ***	(0.004)	-0.008 **	(0.003)
Log total positive income	-0.128 ***	(0.001)	0.010 ***	(0.002)	0.118 ***	(0.001)
Timely filed in past four years	0.592 ***	(0.002)	-0.321 ***	(0.003)	-0.270 ***	(0.003)
Balance due (before remittance)	-0.146 ***	(0.003)	-0.044 ***	(0.005)	0.190 ***	(0.004)
% of income under-withheld	-1.497 ***	(0.010)	-0.153 ***	(0.016)	1.649 ***	(0.015)
50% or more of income not subject to withholding	-0.091 ***	(0.003)	-0.008	(0.005)	0.099 ***	(0.005)
Activity code 266	0.730 ***	(0.065)	-0.187 *	(0.105)	-0.543 ***	(0.093)
Activity code 270	0.246 ***	(0.014)	-0.386 ***	(0.022)	0.140 ***	(0.018)
Activity code 271	0.030	(0.025)	-0.189 ***	(0.041)	0.158 ***	(0.033)
Activity code 272	0.387 ***	(0.013)	-0.143 ***	(0.021)	-0.244 ***	(0.016)
Activity code 273	0.318 ***	(0.013)	-0.218 ***	(0.022)	-0.100 ***	(0.017)
Activity code 274	0.080 ***	(0.013)	-0.167 ***	(0.021)	0.087 ***	(0.017)
Activity code 275	0.066 ***	(0.016)	-0.229 ***	(0.027)	0.163 ***	(0.021)
Activity code 276	0.336 ***	(0.042)	-0.238 ***	(0.071)	-0.098 *	(0.051)
Activity code 277	0.210 ***	(0.042)	-0.240 ***	(0.072)	0.030	(0.051)
Activity code 278	0.397 ***	(0.018)	-0.271 ***	(0.030)	-0.126 ***	(0.025)

Activity code 279	0.430 ***	(0.014)	-0.145 ***	(0.024)	-0.285 ***	(0.019)
Activity code 280	0.231 ***	(0.014)	-0.242 ***	(0.024)	0.011	(0.019)
Activity code 266*DIF	-0.003 ***	(0.000)	0.003 ***	(0.000)	0.001 *	(0.000)
Activity code 270*DIF	-0.001 ***	(0.000)	0.000 ***	(0.000)	0.000 ***	(0.000)
Activity code 271*DIF	-0.001 ***	(0.000)	0.000 ***	(0.000)	0.001 ***	(0.000)
Activity code 272*DIF	-0.001 ***	(0.000)	-0.000 ***	(0.000)	0.001 ***	(0.000)
Activity code 273*DIF	-0.001 ***	(0.000)	0.000 ***	(0.000)	0.001 ***	(0.000)
Activity code 274*DIF	-0.001 ***	(0.000)	0.000 ***	(0.000)	0.001 ***	(0.000)
Activity code 275*DIF	-0.001 ***	(0.000)	0.000 ***	(0.000)	0.001 ***	(0.000)
Activity code 276*DIF	-0.001 ***	(0.000)	0.000	(0.000)	0.001 ***	(0.000)
Activity code 277*DIF	-0.001 ***	(0.000)	0.000	(0.000)	0.001 ***	(0.000)
Activity code 278*DIF	-0.000	(0.000)	-0.000 **	(0.000)	0.000 ***	(0.000)
Activity code 279*DIF	-0.001 ***	(0.000)	-0.000 ***	(0.000)	0.001 ***	(0.000)
Activity code 280*DIF	-0.000 ***	(0.000)	-0.000	(0.000)	0.000 ***	(0.000)
Year 2012	0.300 ***	(0.004)	-0.077 ***	(0.007)	-0.223 ***	(0.006)
Year 2013	0.727 ***	(0.004)	-0.133 ***	(0.006)	-0.594 ***	(0.006)
Year 2014	1.342 ***	(0.004)	-0.460 ***	(0.006)	-0.882 ***	(0.005)
Year 2015	1.446 ***	(0.004)	-0.536 ***	(0.006)	-0.910 ***	(0.005)
Year 2016	1.534 ***	(0.004)	-0.431 ***	(0.006)	-1.102 ***	(0.005)
Year 2017	1.695 ***	(0.003)	-0.433 ***	(0.005)	-1.263 ***	(0.005)
Year 2018	1.861 ***	(0.003)	-0.511 ***	(0.006)	-1.350 ***	(0.005)
Year 2019	2.063 ***	(0.003)	-0.588 ***	(0.005)	-1.475 ***	(0.005)

† Standard errors in parentheses.

Note: *** indicates significance at the 0.01 level in a two-tailed Z-test

** indicates significance at the 0.05 level in a two-tailed Z-test

* indicates significance at the 0.10 level in a two-tailed Z-test

Table A4. Full Parameter Estimates for Linear Model of Change in Balance Due†

Response Variable: $log(U_{ijt})$

Variable	For P_{ijt}=2 N=3,487,662		
Intercept	5.503	***	(0.017)
ACS weighted average	-0.019	***	(0.006)
CP59 weighted average	-0.009	***	(0.003)
Field collection weighted average	-0.000	*	(0.000)
Married filing jointly	0.068	***	(0.002)
Log total positive income	0.202	***	(0.001)
Timely filed in past four years	-0.177	***	(0.002)
Balance due (before remittance)	-0.185	***	(0.003)
% of income under-withheld	0.337	***	(0.011)
50% or more of income not subject to withholding	0.030	***	(0.003)
Activity code 266	1.111	***	(0.066)
Activity code 270	-0.587	***	(0.014)
Activity code 271	-0.481	***	(0.023)
Activity code 272	-0.827	***	(0.013)
Activity code 273	-1.024	***	(0.014)
Activity code 274	-0.590	***	(0.014)
Activity code 275	-0.654	***	(0.016)
Activity code 276	-0.542	***	(0.033)
Activity code 277	-1.008	***	(0.032)
Activity code 278	-0.944	***	(0.019)
Activity code 279	-0.601	***	(0.015)
Activity code 280	-0.415	***	(0.014)
Activity code 266*DIF	-0.001	***	(0.000)
Activity code 270*DIF	0.001	***	(0.000)

Activity code 271*DIF	0.002 ***	(0.000)
Activity code 272*DIF	0.002 ***	(0.000)
Activity code 273*DIF	0.002 ***	(0.000)
Activity code 274*DIF	0.002 ***	(0.000)
Activity code 275*DIF	0.001 ***	(0.000)
Activity code 276*DIF	0.001 ***	(0.000)
Activity code 277*DIF	0.002 ***	(0.000)
Activity code 278*DIF	0.002 ***	(0.000)
Activity code 279*DIF	0.001 ***	(0.000)
Activity code 280*DIF	0.001 ***	(0.000)
Year 2012	-0.023 ***	(0.004)
Year 2013	-0.188 ***	(0.004)
Year 2014	-0.091 ***	(0.004)
Year 2015	0.011 ***	(0.004)
Year 2016	0.062 ***	(0.004)
Year 2017	0.085 ***	(0.004)
Year 2018	0.064 ***	(0.004)
Year 2019	-0.010 ***	(0.004)

† Standard errors in parentheses.

Note: *** indicates significance at the 0.01 level in a two-tailed t-test
** indicates significance at the 0.05 level in a two-tailed t-test
* indicates significance at the 0.10 level in a two-tailed t-test