



2021 IRS/TPC Research Conference

Abstracts of Papers with Bios of Presenters

Session 1: Improving Individual Taxpayer Compliance

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

Ellen Badgley, Kyle Furlong, Lucia Lykke, Abby Ng, and Leigh Nicholl (MITRE); Alan Plumley (IRS, RAAS)

This paper addresses asks the question: when there are one or more audits of individuals in a tax preparer network do the unaudited taxpayers in that network differ in their subsequent tax reporting compared to taxpayers in a preparer network in which there were no audits?

We use data from Tax Years 2006-2018 for one category of correspondence audit that focuses on individuals claiming sole proprietor expenses. We highlight two facets that we hypothesize are central to estimating the general indirect effect in the context of tax preparer networks. First, we use operational audit selection criteria to define a population that would be *eligible* for this type of correspondence audit but who were *not audited*. Among these, our treatment group includes taxpayers who **share a tax preparer with taxpayers who were audited for this category of audit** and our control group has **no audits of this category in their tax preparer network**. This approach takes a narrow focus in order to account for selection bias related to operational audit procedures. Second, we explore the assumption that in order for taxpayers to be affected by another's audit, they must remain within that tax preparer's network long enough for that audit to be known to the tax preparer. That is, the validity of a study on this topic rests on the ability to observe and measure churn within tax preparer networks over time. We present status transition flows to understand how long taxpayers in the treatment and control groups remain with their tax preparers, and discuss the implications for defining a population to model the general indirect effect.

We implement these approaches in a preliminary predictive model to estimate the indirect effect of audits propagated through tax preparer networks. We show results from linear mixed effects models that compare total tax reporting over time between taxpayers in an audited network and taxpayers with no audits in their network. We also discuss next steps for expanding the scope beyond a single category of correspondence audit.

Kyle Furlong is a senior computational social scientist at The MITRE Corporation. He holds a master's degree in Public Policy with an emphasis in Data Analytics from Carnegie Mellon University and a bachelor's degree in Mathematics and Music from Gettysburg College.

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

Paul R. Organ and Joel Slemrod (Univ. of Michigan); Alex Ruda and Alex Turk (IRS, RAAS)

According to the canonical model, tax evasion is constrained by the threat of detection and punishment. Traditionally, the punishment for actions deemed to be evasion has been largely financial and, very occasionally, imprisonment. But there is a third tool, known as collateral sanctions—measures that rescind government-provided benefits or privileges, and are usually enforced by an agency other than the tax agency. Although empirical analyses of the effectiveness of traditional enforcement policies have proliferated, little attention has been paid to the effectiveness of collateral sanctions. We address that gap by examining the recent introduction of a U.S. initiative that restricts passport access to taxpayers with substantial tax debt (referred to as passport “certification”). Focusing first on the subset of certified taxpayers who were denied a passport-related request, we examine behavior before and after request denial, compared to a control group, to estimate the direct effect of the denial. We then take advantage of a randomized controlled trial (RCT) during the rollout of this new authority to identify the effect of certification on the larger set of taxpayers subject to passport certification. We also consider the potential for additional indirect effects of this new authority. We find that denied passport requests have an immediate and strong positive effect on compliance actions for many of those denied a passport request. We find smaller but non-

trivial effects of certification overall, consistent with our expectations that the importance of passport certification will vary considerably across taxpayers.

Paul Organ is a Ph.D. Candidate in Economics and Business at the University of Michigan, and a volunteer with the IRS through the Joint Statistical Research Program.

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

Emily Y. Lin, Ankur Patel, and Alexander Yuskavage (Office of Tax Analysis, U.S. Department of the Treasury)

Because the Earned Income Tax Credit (EITC) is the largest tax benefit program for workers, compliance with EITC rules is of particular interest to policymakers. But despite frequent changes and adjustments to EITC statutes, regulations, and tax administration strategies since 2002, the EITC error rate has remained relatively stable. This stability is despite dramatic shifts in the role played by third party tax preparers, who have historically acted as an important intermediary in filing EITC claims. This paper attempts to understand the role that tax preparers play in EITC compliance by constructing a panel of taxpayers and observing how current EITC claiming decisions depend on past experiences with both the EITC and paid preparers. We find differences between new and experienced claimants, with the current compliance of experienced claimants strongly related to their past compliance. Levels of compliance are further affected by the qualifications of the preparer. From this we conclude that the long-term stability of EITC compliance can be attributed to a mix of consistent levels of compliance by some taxpayers and offsetting changes by others.

Alexander Yuskavage is an economist at the Office of Tax Analysis in the Department of the Treasury. His research primarily focuses on taxpayer compliance and corporate microsimulation. He received his Ph.D. from the University of Wisconsin-Madison in 2010.

Tatiana Homonoff is an Assistant Professor of Economics and Public Policy at NYU's Robert F. Wagner School of Public Service. Her research focuses on identifying areas in which behavioral economics can improve public policy, primarily in the areas of tax policy, program participation, and consumer finance.

Session 2: Impacts of Variations in Process

Sales Tax Administration and the Real Economy

Roger M. White (Arizona State University)

Research in accounting examines how tax administration affects businesses, but the focus has primarily been on IRS administration of income taxes. We examine how the administration of state and local sales taxes affects businesses. Specifically, we examine whether local government administration of sales taxes (as opposed to state government administration) affects the real economy. We study this question in the Florida hotel industry, as counties in Florida can choose to locally administer the county-level tourist tax or have state revenue officers administer the county-level tax in conjunction with the state-level tourism tax. Local administration is popular because it supports local employment (of tax administrators) and provides more stringent enforcement on non-commercial operators subject to the tax (AirBnB, vacation homes rented by owners, etc.). State administration, however, has fewer compliance costs for hotels. For example, managers need only file a single tourist tax return per month (as opposed to one for each county). Commentary from the profession suggests these compliance costs can be considerable, and that local administration of taxes increases tax complexity and associated compliance costs. We find evidence supporting these claims, as our models using county and year fixed effects document that local administration of tourist taxes predicts lower total hotel payroll, fewer hotel employees, and fewer hotels at the county-year level.

Roger White is an assistant professor of accounting at Arizona State University. His research focuses on regulation and fraud in markets.

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

Deandra Reinhart, Rafael Dacal, and Ariel Wooten (IRS, SB/SE); **Jonathan Curtiss** (MITRE)

Many administrative processes conducted by the IRS Small Business/Self-Employed (SB/SE) Division are complex and include a dynamic system of manual and automated processes. To determine how potential process

changes may impact the overall system and measures, SB/SE could conduct a pilot to measure impact. However, these pilots typically result in “down time,” as employees participating in the pilot are trained on the tested process. A discrete event simulation model is an alternative; it is a simple, yet versatile, tool to study problems and conduct experiments relating to resource allocation, scheduling, and capacity planning for systems that drive filing and payment compliance activities, with minor disruption to existing operations and systems. With the assistance of the MITRE Corporation, SB/SE Research developed a proof-of-concept simulation model of the Automated Underreporter (AUR) process. AUR is a complex automated process that compares information returns sent to the IRS by third parties with taxpayers’ filed tax returns to identify underreported income and/or overstated deductions. The AUR process also includes mailing notices to taxpayers, answering incoming phone calls, and receiving and processing correspondence. This system is a high-volume process that can be chronologically sequenced, lending itself well to discrete event simulation. As a result of this project, SB/SE Research found that the IRS can use discrete event simulation models to gain insights leading to process improvements, to identify and remove bottlenecks, and to project the impact of changes to work volume, processing, and staffing without impacting existing work processes. This proof-of-concept project focused on the AUR Philadelphia Campus process.

The primary purpose of the AUR discrete event simulation model is to design and test potential experiments to inform decisions. The simulation model enables modification of process assumptions, adjustment of minor processing steps, and shifting/adding resources, enabling stakeholders to evaluate the impact of changes on the process deliverables and measures. Hence, this modeling method allows IRS operations to obtain data driven recommendations without impacting tax administration.

Rafael Dacal is a Senior Operations Research Analyst for the Small Business/Self-Employee (SB/SE) Division of IRS. He received a Bachelor of Arts and Masters of Arts in Economics, and a Masters of Science in Finance from Florida International University and a Doctor in Business Administration from Georgia State University.

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

Derrick Dennis, Jennifer Ferris, **Chloe Gagin**, Tuba Ozer-Gurbuz, Julia Shiller, and Christopher Williams (IRS, SOI)

This study sought to measure the impact that amended returns would have on SOI tax return statistics. Based on the results, there is an important observation about SOI estimates that can be made for individual and corporate income returns. In this preliminary study looking at 2013, some 2.3 percent of individual income tax returns (Forms 1040) and 0.52 percent of corporate income tax returns (Forms 1120) were amended (using Forms 1040-X and 1120-X). These amendments decreased the estimated total tax from individual filers by only \$2.1 billion (0.16 percent) and decreased the estimated total tax from corporate filers even less, by \$87.8 million (0.03 percent). Thus, the study’s review of individual and corporate tax returns from 2013 suggests that while SOI tax return statistics would be affected by post-filing adjustments, current SOI statistics do not significantly change.

Gloria “Chloe” Gagin is an economist at the IRS’ Statistics of Income Division in the Corporation, Partnership, & International Branch. Chloe earned an M.A. in economics in 2018 from American University.

Jennifer Stratton is a senior economist at the Government Accountability Office, specializing in tax policy. Her recent work includes examining tax-exempt hospitals, international dispute resolution, and transfer pricing. Previously she worked at the Congressional Budget Office, where she specialized in international and corporate taxation. She holds a B.S. in economics from George Mason University, an M.A. in economics from Johns Hopkins University, and a Ph.D. in economics from George Washington University.

Session 3: Developments in Technology and Analytics

New Approaches to Estimating the Extent of Nonfiling

Tom Hertz, Pat Langetieg, Mark Payne, and **Alan Plumley** (IRS, RAAS); Maggie R. Jones (U.S. Census Bureau)

The Internal Revenue Service (IRS) has been estimating the extent of individual income tax nonfiling for many years—both in terms of the number of missing returns and the unpaid tax. The IRS has some information from third parties about nonfilers’ income, but without a tax return, it lacks current information about the filing status and tax benefits (e.g., deductions and credits) to which a nonfiler may be entitled. An individual’s filing status (e.g., single, married filing jointly, etc.) is needed to determine whether a tax return is required, and all of that missing information is needed for determining tax liability. These must either be estimated or obtained from other

sources. The natural source of such information for individuals is surveys conducted by the Census Bureau—specifically, the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC). The IRS has used this survey in several alternative approaches to estimate the extent of nonfiling, but each of these methods has significant weaknesses. Most of those weaknesses can be overcome, however, using new tax administrative data that can be linked to the survey data through the cooperative efforts of the IRS and the Census Bureau. This new capability allows us to match the CPS-ASEC data to a full set of IRS administrative data, anonymously, for most CPS-ASEC respondents—both filers and nonfilers. This allows us to estimate the tax obligations and payment history of each linked individual in the CPS-ASEC survey with much greater certainty than previously possible, yielding improved estimates of the extent of nonfiling. This paper will provide an overview of this new methodology and explain how we have projected the linked survey and tax records to the entire population, taking into account the fact that not all CPS-ASEC records can be matched to IRS administrative records. We conclude by incorporating similar tabulations of late filers from IRS administrative records to derive a comprehensive updated estimate of the Tax Year 2010 individual income tax nonfiler tax gap.

Alan Plumley is a Technical Advisor in the IRS Office of Research, Applied Analytics, and Statistics. His expertise is in the areas of compliance measurement and modeling for efficient workload selection and resource allocation. A 36-year veteran of IRS research, Alan earned his Ph.D. in public policy from Harvard University.

Using Uplift Modeling to Improve ACS Case Selection and Compliance Outcomes

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

In 2016, as part of the Servicewide Future State Initiative, the IRS launched the Automated Collection System (ACS) Optimization and Notice Redesign initiative. The goal of this initiative is to streamline and accelerate the collection of tax debt, and specifically to improve performance in three key areas:

1. **Improve Taxpayer Experience** by directing taxpayers towards self-service tools such as Online Payment Agreement (OPA) and Direct Pay;
2. **Increase Taxpayer Compliance** by designing notices that encourage taxpayers to pay their balances due through full payments, partial payments, or installment agreements; and
3. **Reduce IRS Costs** by decreasing the volume of phone calls and correspondence the IRS receives.

Currently, ACS collection notices—such as the LT11 and LT16—are issued only to a subset of eligible taxpayers, due to resource constraints. Thus, optimizing ACS performance involves not only improvements in **notice design** (i.e., increasing the effectiveness of the notice itself), but also improvements in **case selection** (i.e., deciding which taxpayers should receive each notice).

In the four years since launching the ACS Optimization and Notice Redesign initiative, the IRS has conducted a series of randomized control trials to test the effectiveness of alternative notice designs and case selection methods. In the first test, conducted in 2017, six redesigned LT16 notice designs were tested against the existing LT16, which was used as a control. The redesigned LT16 notices included changes to wording and format that, collectively, guided taxpayers towards desired behaviors and away from undesired behaviors.

In addition to testing the impact of redesigned notices, RAAS and SB/SE have explored the use of “uplift models” to optimize case selection when the IRS is unable to treat all eligible cases. These models estimate the impact of a proposed treatment on each potential case, so that decision makers can then decide how to allocate treatment resources. In this paper, we present findings from an uplift modeling test using data generated by the 2019 LT11 Notice Redesign Pilot. Uplift models for dollars recovered and incoming phone calls were used to predict the potential treatment effects for taxpayers in ACS if they were selected to receive an LT11 notice. These uplift predictions will be used to implement a case selection methodology based on the resource constrained optimization of uplift in dollars collected subject to a cost constraint primarily driven by uplift in incoming phone calls.

Building on the lessons learned from prior pilots, in this iteration predictive models were built to specifically identify taxpayers who are expected to receive a refund offset in the near future, and ensure their cases are not selected for receipt of a notice. A key goal of this effort was to predict latent, underlying features of taxpayers that are likely to have a large impact on their behavior, primarily whether they will have a refund offset.

Jan Millard is a Program Manager in the IRS Office of Research, Applied Analytics, and Statistics (RAAS). Jan has worked for the IRS for over 36 years. This knowledge and expertise have proven effective in applying behavioral insights to notice design and innovative ideas toward case selection within the IRS collection population. Jan holds a Master of Music degree from the Conservatory of Music, University of Missouri Kansas City (UMKC).

Recent IRS Discriminant Function (DIF) Model Improvements

Getaneh Yismaw, Taukir Hussain, Drew Johns, Jon Creem, and Mary-Helen Risler (IRS, RAAS)

The IRS uses Discriminant Function (DIF) score to rank returns by the likelihood of a significant tax change. DIF is one of the main audit selection tools for IRS field examinations. The supervised machine learning model behind the DIF scores is developed using data from the IRS National Research Program (NRP), comprising examination results from a risk-based stratified, random sample representative of filed individual income tax returns. Basing DIF model development on NRP data allows us to develop statistically unbiased and objective risk models. Beginning with Tax Year (TY) 2006, the NRP has been conducting annual studies of individual income tax (Form 1040) reporting compliance. The annual sample data allows us to adopt a dynamic model development and implementation strategy—resulting in significant increases in dollars per case as well as reduced selection of unproductive returns.

IRS groups domestic individual income tax returns into twelve mutually exclusive groups (activity codes) based on return characteristics. The activity codes are used for examination planning, selection and tracking. Using NRP data from TY 2006 onwards, we put into production improved and updated F1040 DIF models over four rounds of development. Since calendar year 2013, all twelve activity codes have been updated at least once, including five activity codes that have been updated twice. The last round of updates resulted in the implementation, beginning in 2019, of new models for 5 activities codes. Moreover, another six updated and improved models are being implemented for production in 2022.

This analysis demonstrates the value of NRP data, and the significant DIF model improvements made by analyzing the outcome of returns examined after the first three rounds of DIF model updates for five activity codes.

Getaneh Yismaw is a Senior Operations Research Analyst within the IRS Knowledge Development and Application Division, RAAS. Since joining the IRS in late 2008 he has primarily been involved with research related to predictive workload selection modeling. He has a B.Sc. Degree in Statistics from Addis Ababa University (Ethiopia), M.S. in Mathematical Statistics from University of Texas at Dallas (UTD), and was a Ph.D. candidate at Southern Methodist University (SMU).

Brian Erard operates an economics consulting practice in the Washington, DC metro area. He specializes in the areas of tax compliance, enforcement, and administration and has published extensively in academic journals, including the widely cited survey of the economics of tax compliance in the *Journal of Economic Literature*. Brian has consulted on a broad range of tax policy issues with governments and businesses in the U.S., Canada, and abroad. Prior to becoming a full-time consultant, Brian spent a decade in academia, first as an Assistant Professor of Economics at the University of Toronto, and later as an Associate Professor of Economics at Carleton University.

Session 4: Enhancing Taxpayer Customer Experience

Increasing Take-up of the American Opportunity Tax Credit

Crystal Hall (University of Washington), Anne Herlache (IRS, RAAS), and Mary Clair Turner (U.S. General Services Administration)

To promote college access and affordability, education tax credits help with the cost of college by reducing the amount of tax owed. The American Opportunity Tax Credit (AOTC) is partially refundable and the most generous Federal higher education tax credit, with a maximum credit of \$2,500 per eligible student. Yet national estimates find that many eligible college students or their families do not claim the AOTC.

We designed and implemented a multimodal, coordinated communications intervention aimed at increasing AOTC take-up. To increase the likelihood that eligible taxpayers open and read the materials, the communications

bundle varied the mode of communications (emails and letter), the sender of the communications (University Bursar and IRS), and the target audience (student and primary tax filer). Further, the content of the messages incorporated feedback from student focus groups and insights from the behavioral sciences (e.g., timely reminders, action steps, loss aversion, easy-to-use links to free tax software).

We block randomized students from the University enrolled during the 2019 tax year to a communications bundle treatment group or control group. The bundle included five emails and a letter. We further randomized the letter component. The IRS sent letters to half of the treatment group and the University Bursar sent letters to the other half of the treatment group as a way to explore possible messenger effects.

Preliminary analysis finds a statistically significant effect of the communications bundle on AOTC take-up—increased from 46.9 percent to 48.4 percent (48.6 percent in the IRS letter group and 48.2 percent in the University letter group).

Anne Herlache is a senior analyst with the IRS Office of Research, Applied Analytics, and Statistics, and holds a Ph.D. in social psychology. Her research focuses on the application of behavioral insights in a variety of spaces.

Customer Experience for Small Business and Self-Employed Taxpayers

Kaho Bonhomme, Kathleen Holland, and Janice Hu (IRS, SB/SE)

In response to Executive Order 12862 and the Restructuring and Reform Act of 1998, the IRS has been measuring customer satisfaction after critical interactions and journeys with taxpayers. Customer satisfaction is one result of a series of customer experiences. This research is part of the effort to move from a customer satisfaction framework to a more inclusive customer experience framework. The objective of this paper is to provide an overview and glimpse into the small business and self-employed (SB/SE) taxpayer experience.

OMB Circular A-11, Section 280 defines federal government customer experience as a combination of factors that result from touchpoints between an individual, business, or organization and the Federal Government over the duration of an interaction and relationship. To better understand this experience for SB/SE taxpayers, we reviewed taxpayer profiles, surveys, journey maps, call analysis and focus groups. Our findings include the preferences, expectations, and opinions of SB/SE taxpayers, as well as the high points and challenges these taxpayers face. Results from this research provide the information necessary to make improvements to services and to increase taxpayer satisfaction and compliance.

Kaho Bonhomme is a Management & Program Analyst and a Technical Advisor for SB/SE Research. His research focuses primarily on the customer experience, organizational behavior, and employee engagement. Kaho is an Agile Certified Professional with a Master of Science in Managerial Science & Organizational Change, and a bachelor's degree in Psychology.

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

Yan Sun and Stephanie Needham (IRS, RAAS)

In this paper, we use tax return data to assess the effectiveness of Forms 943 and 944, by examining the tax filings, reporting, and payment compliance associated with these annual employment tax forms. The paper presents evidence that Form 944 filings have been in steady decline over the past decade, and that the Form 944 has not been widely adopted or utilized by small businesses since its introduction in 2006. In addition, even though the original purpose of the Form 944 was to reduce the burden of small businesses, a significant number of large businesses were found to file the Form 944, indicating a high-level of confusion amongst the Form 944 tax filing population. In contrast, the paper finds that the Form 943 filing population has been stable over the last decade, and that there is little confusion in tax reporting and filing amongst this population. Finally, there is strong evidence that Form 944 filers, specifically, large Form 944 filers, have a much higher rate of payment noncompliance, than Form 943 filers.

Yan Sun is a Senior Economist in the IRS Office of Research, Applied Analytics and Statistics (RAAS). Dr. Sun's current research includes tax compliance, fiscal policy, and applied econometrics. She earned her Ph.D. in Economics from George Washington University.

Jacob Goldin is an associate professor of law at Stanford Law School. His research focuses on the taxation of low-income households and the application of behavioral economics to the design of policy. Prior to joining Stanford, Goldin worked in the Office of Tax Policy at the U.S. Treasury Department and clerked for Judge Richard Posner of the Seventh Circuit Court of Appeals. He holds a J.D. from Yale Law School and a Ph.D. in economics from Princeton University.