2020 IRS/TPC Research Conference

Abstracts of Papers with Bios of Presenters

Session 1: Behavioral Responses to Audits

The Specific Deterrence Implications of Increased Reliance on Correspondence Audits

Brian Erard *(B. Erard & Associates), Erich Kirchler and Jerome Olsen (University of Vienna)*

Tax administrations rely on audits as a key tool for promoting and enforcing tax compliance. Since audit resources are costly and scarce, however, they are largely reserved for cases with substantive compliance risks. Largely in response to budgetary pressures and shifting priorities, there has been a seismic shift in the composition of IRS audits over the past three decades. Whereas face-to-face audits accounted for the majority (62 percent) of all examinations of returns filed in 1990, the lion's share (81 percent) of all audits of returns filed in 2017 were conducted through correspondence. In comparison with face-to-face audits, correspondence examinations tend to be more narrowly focused and less costly to undertake. At the same time, they are more impersonal. In this study, we investigate the implications of the growing reliance on audits by mail for future taxpayer reporting behavior. Our study relies on a large and unique data base that includes audit and comparison samples of self-employed taxpayers covering two different tax years: 2010 and 2014. To account for non-random audit selection, we employ an inverse probability weighting estimation methodology. The results indicate that face-to-face audits are consistently effective in promoting future reporting compliance. On the other hand, the impact of correspondence audits is more nuanced. Correspondence audits that are undertaken shortly after filing (prior to the filing for the subsequent tax year) tend to have a counter-deterrent effect, reducing reported taxes by 6 to 15 over the two years following the examination. In contrast, correspondence audits that take place later (after the next year's tax return has been filed) have a pro-deterrent effect, similar in size to that observed for face-to-face examinations. We suspect that these contrary outcomes may reflect differences in the types of issues or taxpayers that are addressed over the correspondence audit cycle. However, more research is needed to understand the reasons underlying this result. More generally, the findings point to a need for further investigation into the proper balance between face-to-face and correspondence examinations.

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.*

The Specific Indirect Effect of Correspondence Audits: Moving from Research to Operational Application

Leigh Nicholl*, Maxwell McGill, Lucia Lykke (MITRE Corporation), Alan Plumley (IRS, RAAS)*

The specific indirect effect of a tax enforcement activity, such as an audit, refers to the subsequent compliance behavior of the audited taxpayer. That is, do audited taxpayers change their compliance behavior after an audit compared to similar taxpayers who were not audited? In this study, we build on prior research to address this question, and further, we consider how this knowledge can be applied to improve IRS audit planning activities to maximize the return on investment for different types of correspondence audit.

First, we do not treat all categories of correspondence audit the same. Quantifying the specific indirect effect of an operational audit (as opposed to a randomly selected research audit) is notoriously difficult because we cannot observe what an audited taxpayer would have done in the absence of the audit. In this study, we address this challenge by aligning our analyses (including how we define our "treatment" group of audited taxpayers and our "control" group of not audited, but similar taxpayers) with current operational procedures and definitions used by the IRS's correspondence audit program. We exploit knowledge of correspondence audit eligibility and

prioritization criteria for five distinct categories of correspondence audit, each with unique operational criteria, to minimize selection bias (i.e., to ensure that the "control" group of unaudited returns is as much like the group of audited returns as possible) and to align our findings to operations. Second, we translate model estimates for into actual dollar values representing the estimated difference in total tax paid between audited and not audited taxpayers after audit or audit eligibility. The objective of this activity is to generate estimates of the indirect effect that can be added to, or compared with, dollar value estimates of direct revenue generated from audits. This estimate is crucial for translating research on indirect effects into operational decision-making, such as how to efficiently allocate the budget to maximize the combined direct and indirect effects of the audits. Results show that the specific indirect effect of audit, as measured by the difference in estimated total tax paid in subsequent years, varies across categories of correspondence audit, implying that the IRS should consider these indirect effects when evaluating the return on investment for different types of audit.

Leigh Nicholl *is a statistician at The MITRE Corporation, where she has spent the last two years researching the indirect effects of audits for the IRS. When not studying IRS audits, Leigh's work at MITRE focuses on analyzing healthcare data in collaboration with Boston Children's Hospital to predict and improve patient outcomes. Before coming to MITRE, she was a member of the Duke Biostatistics Core, where she conducted statistical analysis of health record data for Duke's Department of Surgery. She holds a master's degree in Biostatistics from Duke University and a bachelor's degree in Molecular Biology and Math from Colorado College.*

The effect of audit risk and detection risk on tax compliance

James Alm and ***Matthias Kasper*** *(Tulane University)*

This study uses a preregistered laboratory experiment with 333 participants to investigate the effect of tax audits on post-audit tax compliance. An important feature of our experimental design is the addition of audit "efficiency" to our audit mechanism, where "efficiency" is defined as the share of unreported income that the tax agency detects in an audit. This addition allows us to examine the effects of audit efficiency on post-audit reporting. We also study whether enforcement has differential effects on different types of taxpayers, as distinguished by their prior reporting behavior. Contrary to theoretical predictions, we find that tax audits have differential effects on post-audit compliance and that the efficiency of audits determines these responses; that is, while efficient audits increase post-audit tax compliance, inefficient audits have the opposite effect. We find no evidence that this behavior results from a misperception of compound detection lotteries where detection is uncertain. Finally, we find that tax audits (whether efficient or not) increase subsequent compliance of noncompliant taxpayers while they reduce compliance among compliant individuals. However, we find no evidence that tax audits crowd out the intrinsic motivation to comply of honest individuals. Our findings suggest that the specific deterrent effect of tax audits is more ambiguous than much previous analysis suggests, with these effects dependent on the efficiency of the audit process and on the taxpayer's prior reporting behavior.

Matthias Kasper *is a Postdoctoral Fellow at the Department of Economics at Tulane University. He holds a PhD in International Business Taxation from Vienna University of Economics and Business and a PhD in Psychology from the University of Vienna. Prior to joining Tulane he worked as a Postdoctoral Researcher at the Department of Applied Psychology at the University of Vienna. His research investigates determinants of tax compliance behavior. He currently studies how complicated tax laws, enforcement, and the provision of services affect compliance.*

Janet Holtzblatt *is a senior fellow at the Urban-Brookings Tax Policy Center. Over a three-decade career in the federal government, she worked on a broad range of tax policy issues, with emphasis on the tax treatment of families and workers, health reform, the administration of the tax code, and tax simplification. Before joining Urban, Holtzblatt was the unit chief for tax policy studies in the Tax Analysis Division of the Congressional Budget Office. Before that, she was deputy director of the Individual Taxation Division in the US Treasury Office of Tax Analysis and a senior analyst on the Senate Budget Committee. She earned her bachelor's degree in economics and history from the University of Illinois and a doctoral degree in economics from the University of Wisconsin.*

Session 2: New Insights on Taxpayer Behavior

Size, characteristics and distributional effects of income tax evasion in Italy

Martina Bazzoli (IRVAPP), ***Paolo Di Caro*** *and Marco Manzo (Italian Dept. of Finance), Francesco Figari (Univ. of Insubria, University of Essex), Carlo Fiorio (Univ. of Milan)*

Estimating tax evasion behavior is a daunting task, but the increasing availability of administrative data is providing new tools to the researcher. By building a novel dataset linking the 2013 Italian Household Budget Survey with individual tax records over a period of 7 years, we manage to estimate food expenditure equation disentangling households with prevalent income from self-employment, which is self-declared, from for those with prevalent income from wage and salary, which is third-party reported. The main assumption is that all groups report food expenditure correctly; wage and salary workers report income correctly, whereas self-employed workers underreport their income. The richness and quality of data allows us to reduce the income measurement error to a great extent by: (I) using administrative data instead of survey data; (ii) using a 7-year-average of current income as permanent income instead of current income, (iii) using stock of wealth data as an instrumental variable for the remaining transitory fluctuations in the average income. And (iv), to depart from the usual constant share of underreporting, showing that underreporting heterogeneity among self-employed is significant, being larger for singles and smaller for .households with secondary-school educated heads. The estimated heterogeneous underreporting behavior of households combined with the use of the Italian module of EUROMOD, a tax-benefit microsimulation model, allowed an estimation of the distributional effects of income underreporting, suggesting that tax evasion accounts up to 28 billion euro revenue loss, with almost 60% of the missing revenue.

Paolo Di Caro *is a tax economist at the Italian Ministry of Economy and Finance (Department of Finance). He has been representing the Department of Finance at different international, European, and national working groups and commissions on the study of tax evasion, including the OECD Advanced Tax Gap Analysis Community of Practice, the European Commission Tax Gap Project Group, the EU Code of Conduct of Business Taxation Group, and the Italian Commission on the Study of Tax Evasion. Prior to join the Ministry of Economy and Finance he worked as researcher at the University of Catania (2014-16), Ernst & Young Public Sector Division (2010-2016), and Harvard University (2012-2013).*

The Effects of an Employment Tax Enforcement Regime on US Small Business and Proprietor Payment Compliance

Rafael Dacal *(IRS, SB/SE)*

This study attempts to identify ways to improve voluntary compliance and minimize taxpayer burden, but also tries to understand the behavior of taxpayers' compliance given the compliance regime. The study explores how safe harbor provisions may change the behavior of employment tax payment for small business and self-employed taxpayers. The question is if a safe harbor provision can reduce the proclivity of authorized individuals to implement a payroll tax dilemma strategy and whether the regime can reduce payment noncompliance in time of economic distress. An online experiment was conducted with individuals from small and self-employed firms. The experiment followed a similar design to Alm et al. (1992), using scenarios to examine options and make choices in an online setting (instead of laboratory) with the hope of mimicking a real-life scenario and require multiple rounds. During the experiment subjects were asked to voluntarily report and pay employment taxes. Once filing and payment of employment tax was completed, participants were subject to a random audit. If underreporting and/or underpayment were detected, the participant was required to pay a penalty as prescribed by the US tax code. After two initial rounds, subjects were asked if they wish to participate in a voluntary safe harbor program. After several rounds under a safe harbor provision, an income shock (i.e., lower business revenues) was introduced to test the viability of an employment tax safe harbor provision on payment compliance. A total of 205 individual participated in the study. Analysis of variance (ANOVA), nonparametric test, and logistic regression models were used to further analyze the data. This study found that for subjects who did not subscribed to safe harbor provision but experienced an increased probability of apprehension (i.e. audit) increased their payment compliance. Also, the availability of a safe harbor provision led to many participants to choose the safe harbor provision to avoid audit. The results from this study also suggest that provisions such as a safe harbor can be a method of reducing filing costs, audit costs and ultimately taxpayer burden. On the other hand, the results were inconclusive in determining if such provisions can improve payment compliance. Nevertheless, the outcome of this study can improve timing and accuracy of employment taxes payments and it may improve the accuracy of employment tax payment.

Rafael Dacal *has worked at the Internal Revenue Service for 16 years. He is currently a Lead Operations Research Analyst in the Small Business/Self-Employee Research organization. He served three years in the US Army (2nd Armor and 1st Cavalry Divisions). After his service, he attended college and received a Bachelor of Arts in Economics, a Master of Arts in Economics, a Master of Science in Finance, and a Doctorate in Business Administration.*

Jamie McGuire *is a senior economist at the Congressional Joint Committee on Taxation. His areas of expertise include tax administration and compliance, individual income tax, accounting methods, financial transaction taxes, and tax-exempt organizations. He received his Ph.D. in economics from Stanford University.*

Session 3: Advances in Taxpayer Service

Free Assisted Tax Preparation Outreach Experiments

Rizwan Javaid *and Brenda Schafer (IRS, RAAS), Jacob Goldin (Stanford University), Tatiana Homonoff (New York University), Adam Isen (Department of the Treasury)*

With approximately 4% of taxpayers filing their returns on paper and almost 12% of taxpayers not filing a return at all, there is an opportunity to reach millions of people with information about filing their return using free assistance. Using assistance allows taxpayers to prepare their return electronically and be 32 times less likely to commit math errors than those not using assistance. Additionally, the IRS spends approximately $0.25 processing an electronic return as opposed to over $4.00 for a paper return.

In Filing Season 2017, the IRS conducted a postcard outreach experiment with prior paper filers to encourage the use of free assisted tax preparation methods. Information about in-person Voluntary Income Tax Assistance (VITA) sites and free online assistance, via the Free File program, was provided in the outreach. The outreach resulted in a significant increase in the use of VITA, particularly when addresses to the nearest VITA sites were provided to the taxpayer. A follow up study was designed for Filing Season 2019 using these results.

For the Filing Season 2019 outreach, the IRS sent a random sample of taxpayers, both prior paper filers and prior nonfilers, letters with information about VITA, Free File, or both programs. In this experiment, all treatments with VITA information included addresses to the nearest two sites to the taxpayer. For communications that included information on Free File, half of the treatments provided the general information Free File link, while the other half provided a link to the Free File Wizard. The Free File Wizard allows taxpayers to answer questions about their tax situation and then provides a list of free software for which the taxpayer is eligible. The results from this experiment yielded significant increases in overall VITA usage and Free File usage, while significantly decreasing nonfiling.

This paper will discuss the results from the Filing Season 2019 outreach and how filing rates and preparation methods of both groups of paper filers and nonfilers were impacted. We will also discuss how the use of VITA and Free File were impacted by individual treatments and groups of treatments. Additionally, we explore how treatment effects varied by age.

Rizwan Javaid *is an Operations Research Analyst in the Taxpayer Behavior Lab (TBL) of Research, Applied Analytics, and Statistics (RAAS). He has been in TBL since 2016 and with the IRS since 2013. At the IRS, he's worked on several outreach efforts for Collection notices, Automated Substitute for Return (ASFR) notices, non-filer reminder notices, and EITC-related correspondence. In his current position, he continues to work on outreach efforts as well as design and analyze burden surveys to estimate taxpayer burden. Rizwan holds a Bachelor's degree in Industrial and Systems Engineering from the Georgia Institute of Technology and a Master's degree in Engineering Management from Mercer University.*

Enforcement versus Outreach - Impacts on Taxpayer Burden

Anne Herlache, Stacy Orlett, Ishani Roy and ***Alex Turk*** *(IRS, RAAS)*

While IRS has different mechanisms for encouraging taxpayers' compliance with filing returns, it is also concerned about the adverse effect of these mechanisms on taxpayer burden. In this update to a 2019 IRS Research Bulletin paper, we focus on the impact of treatments used in a randomized design on outcomes related to taxpayer burden in two distinct populations of individuals: 1) taxpayers with prior year delinquent returns and 2) taxpayers who filed the previous year but are at risk of not filing their current return.

The experiment comprised of several different treatment paths and a control group. Treatment paths varied on the timing, frequency, and type of contact sent to taxpayers. Over three waves of outreach, taxpayers received a reminder to file, a soft notice, and/or a start in the delinquent return notice process, depending on treatment path assignment. The main outcome presented in the 2019 paper was filing after treatment. In the proposed paper, we expand those results by assessing dollars collected and burden imposed by different treatment paths.

For each treatment, an increase in call volume would show a negative effect on the taxpayer burden, while a reduction in penalty accrued due to noncompliance would indicate a positive impact on the overall burden. We test the change in call volume from control to treatment for each wave of contact. While there is evidence that the treatments are helping taxpayers comply with their filing responsibilities, there is little to no evidence of an increase in call volume from softer types of outreach.

Another aspect of the impact of treatments on the filing behavior is captured through the expected reduction in time to file a return. We perform a survival analysis with time to file as the main outcome and find that some of the treatments significantly reduce the expected time to file. This can be cast as a positive outcome for the taxpayers because reduction in filing time also reduces the penalty accrued while failing to meet compliance standards.

This paper reviews treatment options for addressing nonfiling behavior while also bearing in mind an important consideration for taxpayer-facing IRS actions: taxpayer burden. By viewing the treatment effects in light of the impact on call volume and expected penalties, we provide important insight into the broader context of IRS enforcement versus outreach.

Alex Turk *is the Supervisory Economist for the Policy and Program Impact lab within the Knowledge Development and Application division of RAAS. From 2012 to 2016 he led the Strategic Analysis and Modeling group within the IRS's Collection organization. Prior to that, he served in various capacities within IRS Research from 1995 to 2011. He is also an Adjunct Professor at the University of St. Thomas. Alex received his Ph.D. and BS in Economics from Iowa State University.*

Perspectives on new forms of remote identity proofing and authentication for IRS online services

Rebecca Scollan *and Ronna Ten Brink (MITRE Corporation)*

Registration and authentication to secured IRS applications sets the stage for a taxpayer's experience with IRS online services. Attitudes towards account creation and authentication technology, as well as the technologies' usability, is critical to adoption. We investigated taxpayer perspectives on new forms of remote identity proofing and authentication for IRS online services over two usability studies. In the first, MITRE prototyped a new Secure Access application flow that is compliant with the new NIST Digital Identity requirements and conducted usability testing with tax professionals. In the second, MITRE prototyped a notional credential service provider application flow and usability tested it with taxpayers who have recently owed a balance. We will present findings and recommendations on remote identity proofing and authentication for the IRS context.

Rebecca Scollan *is a Principal Human Factors Engineer at The MITRE Corporation. She researches the user experience of identity proofing and authentication methods for a wide range of audiences and has performed usability research and design across a variety of domains for the IRS for more than 8 years. She earned a M.S. Interaction Design and Information Architecture from the University of Baltimore.*

Mary-Helen Risler *is a supervisory economist in the US Internal Revenue Service's Research, Applied Analytics & Statistics organization. As chief of the Compliance Modeling Lab in the Knowledge Development & Application Division, she oversees a staff of eight technical analysts. The primary focus of the group's analytical work is compliance analysis and estimation-including tax gap estimation-and the development of predictive models for Examination workload selection. She joined the IRS's former Research Division in 1988. As an analyst, her primary areas of compliance research and analysis were employment tax, Earned Income Tax Credit, Automated Underreporter Program, and Affordable Care Act.*

Session 4: Doing More With Less

Can Machine Learning Improve Correspondence Audit Case Selection? Considerations for Algorithm Selection, Validation, and Experimentation

Ben Howard *and Lucia Lykke (MITRE), Alan Plumley (IRS, RAAS)*

The IRS conducts correspondence audits in a wide variety of categories, with each category focusing on potential misreporting on just a few specific lines on tax returns. Typically, returns are prioritized for correspondence audit based on one or more features of the return in question. We demonstrate how machine learning methods that incorporate a broad range of tax return and other characteristics may improve audit case selection (prioritization) through two modeling approaches: 1) find the returns with highest expected assessed and/or

collected revenue that will be generated from the audits; and 2) avoid returns likely to have no-change audit outcomes (that is, audits that result in little or no adjustment). If we could perfectly predict before an audit whether that audit will produce a tax change, and if so, what the revenue/cost ratio will be, then we would rank the returns in declining order of that ratio. Because we do not have perfect knowledge, our pre-audit estimates of audit outcomes are subject to much uncertainty. That means that a modeling approach that seeks to find high-value cases is likely to yield somewhat different results from an approach that seeks to avoid low-value cases.

This study has two parts. First, we discuss the implementation of several iterations of random control trials that compare machine learning models with status quo methods of audit case selection for Tax Years (TYs) 2013-2016 tax returns. Results show that for one type of correspondence audit, the machine learning selection algorithm resulted in a 29 percent increase in revenue in TY2014 compared to the status quo method, while TY2015 results showed little difference between the machine learning algorithm and the status quo selection method. For another category of correspondence audit, the machine learning algorithm did not result in additional revenue, but it did result in a decrease in the no-change rate of approximately 36 percent in TY2013 and a decrease of 25 percent in TY2014 compared to the status quo selection method.

In part 2, we focus on refining our models and research agenda based on lessons learned from the experiments in part 1. For several categories of correspondence audit, we apply a variety of machine learning techniques including regression algorithms, ensembles between classification (seeking to minimize the no-change rate) and regression (seeking to maximize the dollar outcome), and learning-to-rank algorithms. We validate our results during model development using diagnostic visualizations, and we show that strictly minimizing the audit no-change rate may come at the cost of forgoing returns resulting in higher revenue, and vice versa.

Ben Howard *is a senior operations research analyst at the MITRE Corporation. He has been with MITRE for four years, and the majority of that time has been spent working on improving IRS business operations, particularly audit operations. Ben specializes in statistical analysis, machine learning, and analytic pipeline development. He holds a master's degree in Statistical Science and a bachelor's degree in Mathematics with a minor in Computer Science from George Mason University.*

Audit productivity, taxpayer service, and compliance: Can a service mindset overcome a dwindling enforcement budget?

Nina Collum and ***Mary Marshall*** *(Louisiana Tech Univ.), Susan Jurney (Oklahoma City Univ.)*

We examine the very real possibility that audits begin to lose effectiveness, which we label "audit productivity," as tax authorities continue to receive fewer resources. In a simulated compliance environment with real US taxpayers, we predict and find a negative relationship between declining audit productivity and taxpayer compliance in subsequent periods. Although the most obvious solution would be to invest the resources necessary to achieve full audit productivity, this is not feasible in the current budgetary environment. Thus, we propose and test that the inclusion of minimal cost service messages can offset the negative effect of declining audit productivity by influencing how the taxpayers perceive the taxing authority. Specifically, our results show taxpayers who view a service message are more likely to view the Internal Revenue Service as focused on customer service (rather than as focused on punishing criminals). Results contribute to the literature on individual tax compliance behavior, particularly related to the balance between the service and enforcement paradigms.

Mary Marshall *joined Louisiana Tech University as an Assistant Professor of Accountancy in 2018. She earned a Ph.D. from the Darla Moore School of Business at the University of South Carolina in 2018. She also holds a Master of Public Administration from the Hugo Wall School and a Master of Accountancy from the Barton School, both at Wichita State University, Her bachelor's degrees are in Accounting and Management from the Spears School of Business at Oklahoma State University. She has experience working in accounting, primarily focused on tax compliance, and in municipal budgeting. Her research interests include: (1) the judgment and decision making of auditors and tax professionals, with a particular interest in elements of the public accounting firm environment; and (2) the effects of tax fairness and tax policy on taxpayer compliance.*

Using the Internal Revenue Service Program Assessment Model Optimizer to Inform Resource Allocation Decisions

Deandra Reinhart and ***Clay Swanson*** *(IRS, SB/SE)*

As the IRS continues to modernize, it is imperative that workloads and resource allocations are assigned in a logical, data-supported method. SB/SE Research and MITRE have developed a Program Assessment Model (PAM) Optimizer to study current staffing, downstream interactions, and enforcement revenue collected within critical IRS processes, and use this information to optimize new staffing allocations.

The PAM Optimizer is a linear programming model that will allow for informed decisions regarding new workforce allocations and Full Time Equivalent (FTE) needs. Currently the Optimizer is being run with different experimental scenarios which focus around new FTE levels while taking into consideration many constraints such as induced taxpayer responses. We can use the Optimizer to quantify the impact of FTE allocations on enforcement revenue collected and explore alternative resource allocation scenarios.

Clay Swanson *is a senior Operations Research Analyst for SB/SE Research. He has worked for SB/SE Research since 2003, first in St. Paul, MN and currently in Philadelphia, PA. Clay has a Master of Science degree in Mathematical Statistics.*

Arnie Greenland *is Professor of the Practice at the Robert H. Smith School of Business, University of Maryland where he teaches courses in Business Analytics and is principal of a small LLC, AG Analytics. He retired in 2014 from the IBM Corporation, where he was an IBM Distinguished Engineer. His career spanned more than 40 years in which he delivered advanced business analytics services for both public- and private-sector clients, with IRS as one of his primary clients in that period. Arnie focused in his last few years on supporting public agencies engaged in fighting against fraud, waste and abuse, his last major project being technical lead at IBM for the development of IRS's RRP system. He holds numerous patents and authored several professional articles in this area. Arnie is a very active member of the Institute for Operations Research and the Management Sciences (INFORMS). Arnie holds a Ph.D. in Mathematics.*