Analysis Plan

Project Name: Increasing voluntary tax compliance through outreach to clients of return preparers

Project Code: 2007

Date Finalized: 2021-06-22

Project Description

This project aims to improve the outreach and education efforts to return preparers and taxpayers implemented by the Internal Revenue Service (IRS) as part of the agency's broader strategy to reduce errors in claiming benefits by intervening early and ensuring compliance with the law. Specifically, this project aims to build evidence on the effectiveness of the pre-filing season educational letters that the IRS sends to clients of return preparers and the effectiveness of program changes to this strategy. The Letter 6138 informs clients that their return may contain errors in claiming benefits, provides tips to selecting a return preparer, and advises the taxpayer to review the return for errors in claiming certain benefits.

The General Service Administration's (GSA's) Office of Evaluation Sciences (OES), and the IRS's Research, Applied Analytics and Statistics (RAAS) and Wage and Investment (W&I) divisions collaborated on this project to measure the effects of sending the Letter 6138 to clients of return preparers on a number of outcomes, including returns filed with errors claiming certain benefits and clients' tax filing method. Additionally, this project will examine the effects of the Letter 6138 on clients who used the same return preparer, but were not themselves sent a Letter 6138 (i.e., the indirect or network effect of the Letter 6138) and explore whether the Letter 6138 affected the returns prepared by return preparers during filing season 2021. Finally, the project will examine whether sending a modified behavioral insights (BI) Letter 6138 affects the outcomes of clients differently from sending clients the treatment-as-usual (TAU) Letter 6138.

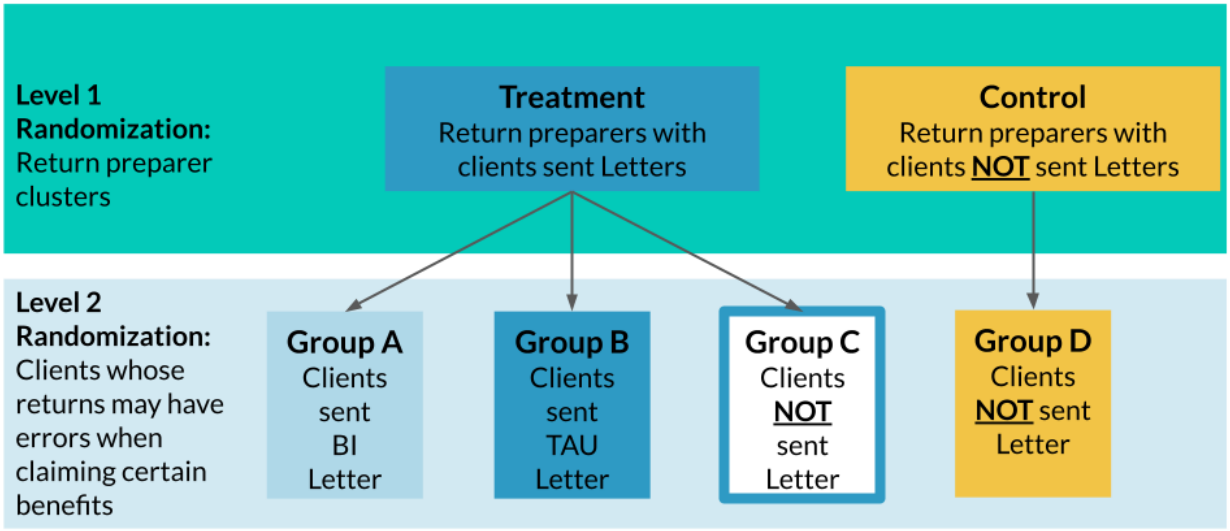
Evaluation Design

This project involves a two-step randomization process which is relevant for both the structure of the data, as well as for the statistical modeling decisions that we made below. In the first step of the randomization process (or the first level of randomization), we implemented a cluster-randomized design where the cluster is defined as a group of clients who used the same return preparer during filing season 2020. In the second step of the randomization process, our sample includes returns filed by the return preparer during the 2020 filing season that may contain errors in claiming certain benefits.[1] In the second step of the randomization process (or the second level of randomization), we randomized these returns (i.e., clients) in treatment return

[1] Note that the return preparers in the sample may have filed additional returns that are not included in the randomization process and are beyond the scope of this analysis.

preparer clusters to be sent the BI Letter 6138, TAU Letter 6138, or not sent any pre-filing season letter. We illustrate the randomization process in Figure 1.

Figure 1: Randomization Process



At the first level of randomization, we grouped together similar return preparers into return preparer blocks and randomized return preparers within these blocks. We describe the sample size by return-preparer-level assignment in Table 1. We aim for each block to have 6 return preparers; however, block size ranges from 3-8 return preparers.

Within these return preparer blocks, we randomized two-thirds of return preparers to have the IRS send some clients the Letter 6138 and one third of return preparers to have no clients sent any pre-filing season letter.

Table 1. Sample size by return-preparer-level assignment

	Treatment	Control	Total
Preparers	1331	663	1994
Clients	52,348	25,344	77,692

At the second level of randomization, we randomize clients, whose return preparers are assigned to the treatment group, in client random assignment blocks. We grouped together clients who used the same return preparer and had similar returns during the 2020 filing season. Within each client block, we randomized approximately 40 percent of clients to be sent a Letter 6138 and approximately 60 percent of clients to be sent no pre-filing season letter. Finally, among clients randomized to be sent a pre-filing season letter, we randomized the version of the Letter 6138 the

IRS sent them. Within each client block, half of clients sent pre-filing season letters were sent the TAU Letter 6138 and half were sent the BI Letter 6138. We summarize the sample size by client-level assignment in Table 2.

Table 2. Sample size by client-level assignment

	Treatment Cluster			Control Cluster	Total
	BI Letter	TAU Letter	No Letter	No Letter	
Clients	10,425	10,424	31,499	25,344	77,692

Glossary of Terms

Before proceeding, a simple glossary of terms is helpful to provide clarification on all terms used throughout this analysis plan:

(1) **Treated Cluster:** a group of clients who used the same return preparer during filing season 2020 who is randomly assigned to have some of their clients sent a Letter 6138;

(2) **Untreated/Control Cluster:** a group of clients who used the same return preparer for filing season 2020 who is randomly selected to have no clients sent a Letter 6138;

(3) **Clients:** clients whose return filed during the 2020 filing season may contain error(s) when claiming certain benefits and who used a return preparer who was randomized as part of this study;

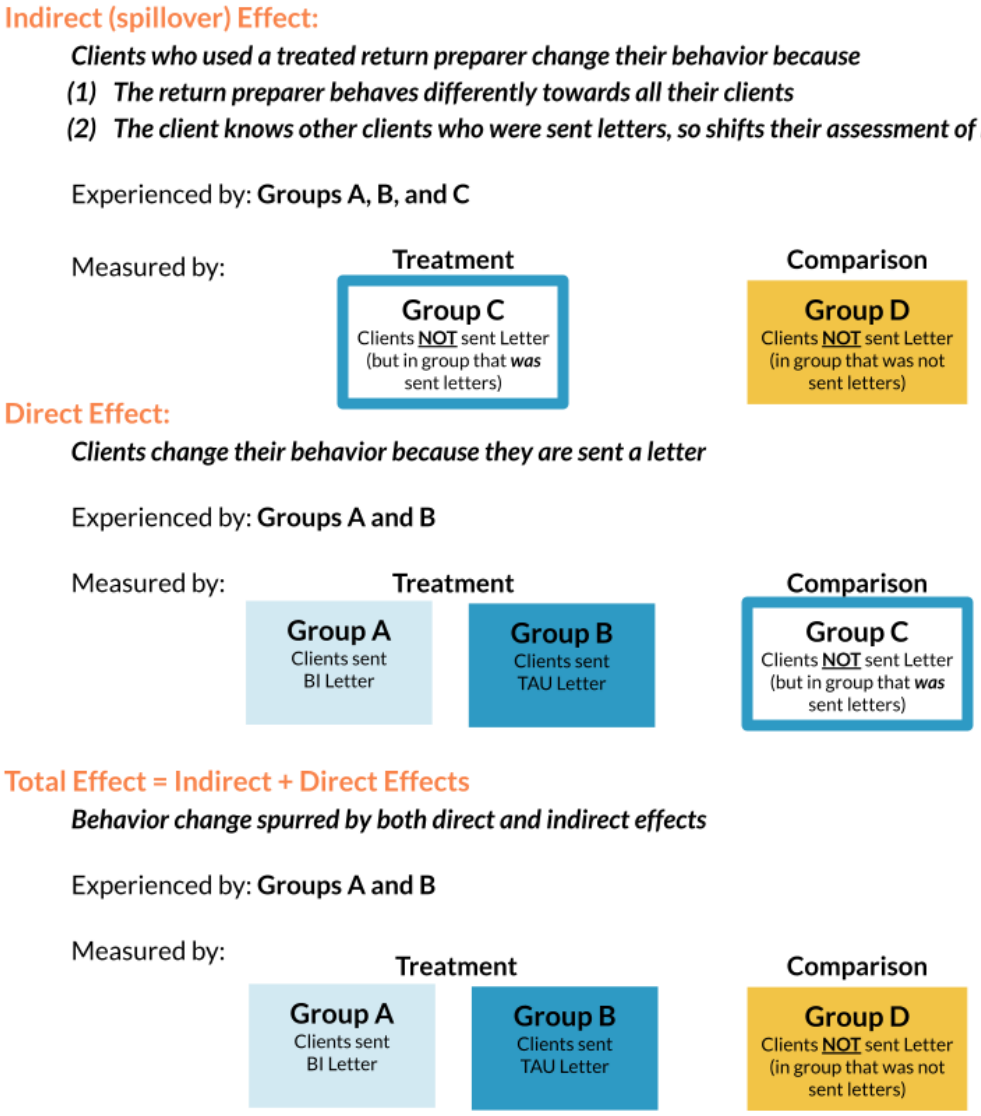
(4) **Indirect Effect (or spillover effect):** the effect on clients of being in a treated cluster—experienced by clients sent or not sent the Letter 6138— where the comparison group to estimate the effect are clients in an untreated cluster

(5) **Direct Effect:** the additional effect (i.e., on top of the indirect effect) of being sent the Letter 6138 among clients in treated clusters—experienced by clients sent the TAU Letter 6138 or BI Letter 6138—where the comparison group to estimate this effect are clients within the same treated cluster who are not sent a pre-filing season letter (i.e., among clients who used the same return preparer during filing season 2020); and

(6) **Total Effect:** the effect on clients of being sent a Letter 6138 and being in a treated cluster where the comparison group is clients in an untreated cluster.

To further clarify the differences between the indirect, direct, and total effects, the following diagram (which also underpins our randomization decisions) may be helpful:

Figure 2. Types of effects and comparison groups



Data and Data Structure

This section describes variables that will be analyzed, as well as changes that will be made to the raw data with respect to data structure and variables.

Data Source(s):
Our primary data source will be processed, return-level data data that W&I pulls for return preparers and clients (i.e., returns) at the end of the 2021 filing season (summer 2021).

The primary time period(s) that will be used for analysis are returns filed during the 2020 filing season (for pre-treatment covariates and blocking) and returns filed during the 2021 filing season (for outcomes). Additionally, data on prior tax years could inform the analysis.

Primary and Secondary Outcome Variables to Be Analyzed:

Primary Outcomes:
We have four different primary outcomes:

- **Change in Filing Method in filing season 2021 (*"Change Method"*):** this variable is binary, and it adopts the value one if a client changed their method of filing their tax return by using a different return preparer (paid return preparer or free tax preparation services) than the return preparer they used during the 2020 filing season, submitted their own tax return, or did not file a tax return.

- **Tax Benefit Error:** this variable is binary, and it adopts the value one if the client files a return that may contain one or more errors in claiming certain benefits, which for the purposes of this study include: the earned income tax credit (EITC), child tax credit/additional tax credit/credit for other dependents (CTC/ACTC/ODC), American opportunity tax credit (AOTC) and head of household (HOH) filing status.[2] It is zero otherwise, including if the client does not claim these benefits or does not file a tax return.

- **Refund Amount:** this variable is numeric, and it reflects the return-level refund amount. For the purposes of this study, this is a measure that will be used to calculate protected revenue, which captures the monetary savings from sending the Letter 6138 to clients.

- **Sum of erroneous dollars:** this is a numeric variable that equals the numeric value for the sum of erroneous dollars that may have been claimed for certain benefits on each return. This is an alternate measure which will be used to calculate protected revenue.

These outcomes are at the return level and include outcome data during the 2021 filing season for the sample of clients (i.e., returns) who were randomized as part of this study. Our analysis focuses on these outcomes to examine the impact of being in a treated return preparer cluster (indirect effect), the impact of being sent a Letter 6138 (or one type of Letter 6138) (direct effect), and the additive impact of being sent a Letter 6138 and being in a treated return preparer cluster (total effect).

Since we measure these client-level outcomes regardless of their filing method during filing season 2021, we do not anticipate any missing data once return processing has completed.

Transformations of Variables:

[2] https://www.eitc.irs.gov/tax-preparer-toolkit/preparer-compliance-focused-and-tiered/compliance

N/A

Imported Variables:

N/A

Transformations of Data Structure:
Thus far, we have referred to clients as the unit of randomization and thus also as the unit of analysis, since it is easier to think of treatment effects acting on people. However, it is technically more correct to refer to the unit of analysis as the tax return, as individuals may file joint returns, or may claim others as dependents on their return, so a "client" in this case may in fact refer to two or more individual people. We conducted randomization by returns, and the data is given at this level, as well.

In most cases, we expect to be able to follow the same people from the 2020 filing season to the 2021 filing season, since it is likely that people who filed jointly in one year will do so again in the next year. However, in cases where a couple has been newly formed or has split (or chosen to file separate returns), we will make the following changes to the data:

> **Individuals who filed separately in the 2020 filings but jointly in the 2021 filing season:** In this case, it is possible that one member of the couple was not in our randomized sample at all, that one person was sent a letter and the other was not (i.e., was in group C or D), or that each individual was sent a different letter. In this case, we will associate the jointly filed return in the 2021 filing season. In other words, the return will be included in the regression once associated with person X and once associated with person Y.
>
> **Individuals who filed jointly in the 2020 filing season but separately in the 2021 filing season:** In this case, we will treat both of the returns filed during the 2021 filing season as being in the same treatment group, but will measure outcomes for each return separately and weight each return accordingly (i.e., at 50 percent).

Since we have no reason to expect differential creation or dissolution of couples across treatment groups, we do not anticipate that these changes to the data will impact our results.

All primary analyses are conducted at the client level (i.e., return level). As described below, in exploratory analyses, we will aggregate client-level data up to a return-preparer-level dataset. To do this, we will restrict the sample to individuals who used one of the return preparers in our sample in the 2021 filing season (i.e., filed using that return preparer in 2021). We will then create variables indicating the total number of returns prepared, and total returns prepared with probable TC errors, by return preparer.

Data Exclusion:
For the purposes of this study, we will exclude client outliers using IRS's typical procedures.

Treatment of Missing Data:

We do not anticipate substantial missing data since our data capture the full sample of taxpayers. We assume that idiosyncratic reasons for not filing tax returns are equal across treatment conditions, and thus the refund amount for these observations will be coded as zero. Not submitting a tax return in filing season 2021 (typically for TY 2020) is an outcome of interest, and thus missing observations are re-coded as a binary variable.

Our analysis will rely on data processed by the end of July 2021. Until then, there may be individuals who have filed their returns, but their return has yet to be processed fully. In this case, outcomes data for some measures will be missing until their return is fully processed.

Descriptive Statistics, Tables, & Graphs

The core figures that we will include in the OES abstract will be four bar charts (using the OES template) for our four primary outcomes of interest.

Primary Hypothesis Tests & Statistical Models

To summarize above, there are three research questions related to this project:

1. What is the effect of sending the Letter 6138 on the behaviors of clients who are sent a letter?
2. What is the effect of sending the Letter 6138 on other clients who are in the same return preparer cluster (but who are not sent a letter)?
3. What is the effect of the different contents of the Letters 6138 on the behaviors of clients who are sent a letter?

The first two research question will be estimated using the following model:

$$y_{cib} = \alpha_0 + \alpha_1 AnyLetter_{cib} + \alpha_2 TreatedCluster_{cib} + P_b + Z_{ci} + \varepsilon_{cib} \qquad (1)$$

Where c indexes for client and i indexes for return preparer in return preparer block b

- $TreatedCluster_{cib}$ = 1 if client c used return preparer *i* (in return preparer block *b)* in the 2020 filing season and return preparer *i* had clients who were sent the Letter 6138, but client *c* was not herself sent a letter (Group C); and 0 if client *c* was sent a letter herself or if return preparer *i* had no clients who were sent Letter 6138 (Groups A, B, or D).
- $AnyLetter_{cib}$ = 1 if client *c* was sent either a BI or TAU Letter 6138 (Groups A or B), and 0 if client *c* was not sent any letter (Groups C or D).
- P_b =a vector of categorical variables used to generate random assignment blocks for return preparers.

- Z_{ci}= a vector of categorical variables used to generate random assignment blocks for clients.
- ε_{cib}= idiosyncratic error term

In equation (1), α_1 captures the total effect of the Letter 6138, whereas α_2 captures the spillover (indirect) effect on clients of either letter. The omitted group from equation (1) is Group D, i.e., clients who were not sent the Letter 6138 and are in a return preparer cluster where other clients were not sent the Letter 6138. We will run this model using OLS with Lin-adjusted covariates, and we will use heteroskedastic robust standard errors (HC2).[3]

To answer research question 1, we test if the total effect, α_1, is statistically significantly different from zero. We will report this as a point estimate with a corresponding p-value. To answer research question 2, we test if the spillover effect (α_1) is statistically significantly different from zero.

To examine differences between the BI Letter 6138 and TAU Letter 6138, we estimate a model of the following form:

$$y_{ci} = \beta_0 + \beta_1 BI_{ci} + \beta_2 TAU_{ci} + \mathbf{\Phi}_i + Z_{ci} + \varepsilon_{ci} \qquad (2)$$

Where *c* indexes for client and i indexes for return preparer in return preparer blocks *b*.

- BI_{ci}= 1 if client *c* who used return preparer *i* in the 2020 filing season was sent a BI Letter 6138 (Group A), and 0 if client c was sent a TAU Letter 6138 or was not sent any letter (Groups B or C).
- TAU_{ci}= 1 if client *c* who used return preparer *i* in the 2020 filing season was sent a TAU Letter 6138 (Group B), and 0 if client *c* was sent a BI Letter 6138 or no letter (Groups A or C).
- $\mathbf{\Phi}_i$=a vector of return preparer fixed effects;
- Z_{ci}= a a vector of categorical variables used to generate random assignment blocks for clients; and
- ε_{ci}= idiosyncratic error term

In equation (2), β_1 captures the direct effect of being sent a BI Letter 6138, whereas β_2 captures the direct effect of being sent a TAU Letter 6138. With the inclusion of return preparer fixed effects,

[3] See Winston Lin. 2013. Agnostic Notes on Regression Adjustment to Experimental Data: Reexamining Freedman's Critique. *The Annals of Applied Statistics* 7(1): 295-318.

the comparison group is clients within Group C, i.e., clients in a treated return preparer cluster who were not sent a Letter 6138. Group D is omitted from this analysis. To test if there is a difference between the TAU and BI Letter 6138s, we conduct an F-test if the difference between these two estimates is statistically significant from zero: $\beta_1 - \beta_2 = 0$.As with equation (1), all models will be estimated using OLS with Lin-adjusted covariates with heteroskedastic robust standard errors (HC2).

In Table 3, we summarize the main hypotheses across these two models and group these tests into distinct families for each outcome.

Table 3: Summary of Tests

Outcome	Test (H_0)	Family
What is the effect of sending the Letter 6138 on the behaviors of clients who are sent a letter? We answer this research question by measuring whether there is a total effect of sending any Letter 6138 on client outcomes.		
Change Method	$\alpha_1 = 0$	1
Tax Benefit Error	$\alpha_1 = 0$	1
Refund Amount ($)	$\alpha_1 = 0$	1
Sum of erroneous dollars ($)	$\alpha_1 = 0$	1
What is the effect of sending the Letter 6138 on other clients who are in the same return preparer cluster (but who are not sent a letter)? We answer this research question by measuring whether there is an indirect effect of sending any Letter 6138 on client outcomes.		
Change Method	$\alpha_2 = 0$	2
Tax Benefit Error	$\alpha_2 = 0$	2
Refund Amount ($)	$\alpha_2 = 0$	2
Sum of erroneous dollars ($)	$\alpha_2 = 0$	2
What is the effect of the different contents of the Letters 6138 on the behaviors of clients who are sent a letter? We answer this research question by measuring whether there is a difference in effects of sending the BI Letter 6138 compared to the sending the TAU Letter 6138 on client outcomes.		

Change Method	$\beta_1 - \beta_2 = 0$	3
Tax Benefit Error	$\beta_1 - \beta_2 = 0$	3
Refund Amount ($)	$\beta_1 - \beta_2 = 0$	3
Sum of erroneous dollars ($)	$\beta_1 - \beta_2 = 0$	3

Inference Criteria, Including Any Adjustments for Multiple Comparisons:

We will apply multiple hypothesis corrections within each of the three families of tests, each with four tests, described in Table 3. Because some of the outcomes within a family may be highly correlated, we will run simulations to control the family-wise error rate, in line with #7 in Alex Coppock's guide.[4] We will use a cutoff of p = 0.05 to determine statistical significance (with stars according to + p = 0.10, * p = 0.05, ** p = 0.01). All tests will be two-tailed.

Limitations:

The following include reasons why we may expect effects during this filing year to be unique:

- If a taxpayer earned less in 2020 compared to 2019, they have the option to choose which year to rely on when calculating EITC or ACTC credits.[5] [6]
- The Economic Impact Payment can be claimed as a tax credit if it was not received during the year.
- The residual impact of the COVID-19 Pandemic could also influence return preparer behaviors in many ways. For example, some return preparers may not have been agile enough to switch their business models to work remotely, and return preparers (as well as clients) may have been impacted from a health standpoint.
- Fewer VITA sites are operating during the 2020 filing season, such that it may be more difficult to substitute away from a paid return preparer.
- Up to $10,200 in unemployment insurance benefits are exempt from taxes. This and other large changes to income for many individuals in the past year may change the types of people who can claim certain tax benefits this year, as well as change the client base of the return preparers.
- The 2020 filing season started and ended later than usual (returns were not accepted until February 12, but can be submitted as late as May 17, 2021), potentially changing the number of clients that a return preparer sees during the filing season.

Exploratory Analysis:

[4] https://egap.org/resource/10-things-to-know-about-multiple-comparisons/
[5] https://www.eitcoutreach.org/blog/new-lookback-rule-may-help-you-qualify-for-a-larger-tax-refund/.
[6] This could have led to returns also getting stopped in the IRS' Error Resolution System.

A) Additional Outcomes

Using equation (1) above, we are also interested in examining the following client-level outcomes. These outcomes provide insights into the margins upon which client behavior may have changed due to the Letter 6138.

1. **Earned Income Tax Credit (EITC) Error**: is a binary variable that adopts the value one if the return may contain an error when claiming EITC.

2. **American Opportunity Tax Credit (AOTC) Error**: this variable is binary, and it adopts the value one if the return may contain an error when claiming AOTC.

3. **Combined Child Tax Credit Error:** this variable is binary, and it adopts the value one if the return may contain an error when claiming ACTC/CTC/ODC.

4. **Review Return**: this is a binary variable that adopts the value one if the client reviewed their return from filing season 2020 and filed a 1040X modification to correct any errors in that return..

5. **Self-File**: this is a binary variable that adopts the value one if the client filed their tax return during the 2021 filing season themselves, and 0 otherwise.

6. **Did Not File:** this is a binary variable that adopts the value one if the client did not file their tax return during the 2021 filing season.

Because the refund amount and the sum of erroneous dollars are key variables that the IRS uses to calculate revenue protected, and revenue protected is typically calculated as a difference-in-differences, as part of our exploratory analyses, we will also run the following specifications for those two primary outcome variables:

$$y_{cibt} = \alpha_0 + \alpha_1 AnyLetter_{cib} + \alpha_2 TreatedCluster_{cib} + \alpha_3 Post_t + \alpha_4 Post_t * AnyLetter_{cib} + \alpha_5 Post_t * TreatedCluster_{cib} + P_b + Z_{ci} + \varepsilon_{cib} \quad (1b)$$

$$y_{ci} = \beta_0 + \beta_1 BI_{ci} + \beta_2 TAU_{ci} + \beta_3 Post_t + \beta_4 (Post_t * BI_{ci}) + \beta_5 (Post_t * TAU_{ci}) + \mathbf{\Phi}_i + Z_{ci} + \varepsilon_{ci} \quad (2b)$$

In the above equations, *Post* is a variable equal to 1 if the outcome comes from tax season 2021, 0 if it comes from tax season 2020. All other variables are defined as in equations (1) and (2) above. The relevant coefficients are α_4,α_5, β_4, and β_5, which represent the treatment effects after subtracting off the baseline differences between the treated and the control groups.

B) Letter Type Specific Effects:

In exploratory analyses, we will test if there is a direct effect of the TAU Letter 6138 or the BI Letter 6138 on the outcomes of interest. To do this, we will use the results from equation (2) and test if $\beta_1 = 0$ (Null hypothesis), which will test if BI Letter 6138 had an effect compared to clients in treated clusters who were not sent letters (Group C), and $\beta_2 = 0$, which will test if the TAU Letter 6138 had an effect compared to clients in treated clusters who were not sent letters (Group C). We may also compare the effects of each letter to the Group D, i.e., clients who were not sent a Letter 6138 in untreated return preparer clusters. To do this, we will estimate a regression of the following form:

(3) $$y_{cib} = \delta_0 + \delta_1 TreatedCluster_{cib} + \delta_2 BI_{cib} + \delta_3 TAU_{cib} + P_b + Z_c + \varepsilon_{cib}$$

Where BI_{cib} adopts the value 1 if client c was sent a BI Letter 6138 and TAU_{cib} adopts the value one if client c was sent a TAU Letter 6138. All other terms are defined above in equation (1). To test if there is a difference between the behavior of clients sent the BI Letter 6138 compared to clients in Group D, we test if $\delta_2 = 0$. To estimate the effect of the TAU letter relative to group D, we test if $\delta_3 = 0$.

C) Return-preparer-level Analysis:

We are interested in the impacts on return preparer behavior during the 2021 filing season of having some of their clients sent Letter 6138. Return preparer responses could include choosing not to prepare returns at all (dropping out of the market) during the 2021 filing season, or, contingent on remaining in the market, changing their filing behavior.

If return preparers remain in the market, the returns they prepare in the 2021 filing season will include (1) returns they prepare for returning clients who were part of client-level analysis for this study, (2) returns they prepare for returning clients who were not part of the client-level analysis for this study, and (3) returns they prepare for new clients. Thus outcome variables in this section are aggregates of measures of these clients. Note this *excludes* clients who used the return preparer during the 2020 filing season but did not use the return preparer during the 2021 filing season (even if these clients were randomized as part of this study), as these clients would not have been impacted by any changes to the return preparer's behavior.

As a result, return-preparer-level impacts include compositional effects of the Letter 6138 on the clients that a return preparer serves and changes to return preparers' behaviors when filling out returns for their clients. We cannot separate these changes. Note that this family of outcomes is exploratory, so we do not plan to make adjustments for multiple hypotheses for this section, and we may not report on these outcomes.

To more comprehensively examine the impact on return preparer behaviors of having some of their clients sent a pre-filing season letter, we will conduct several analyses at the return preparer level. Specifically, we are interested in answering the following questions:

(1) What is the effect of having the IRS send some clients the Letter 6138 on the likelihood that a return preparer stops preparing tax returns during the 2021 filing season?

(2) What is the effect of having the IRS send some clients the Letter 6138 have on the number of returns a return preparer prepares (which is zero if the return preparer has dropped out; we will run this based on the actual number of returns prepared, and will not bottom-code the number of returns for any return preparer)?

(3) What is the effect of having the IRS send some clients the Letter 6138 have on the number of returns a return preparer files for their client that may contain errors when claiming certain benefits?

(4) What effect did having clients sent Letter 6138 have on the proportion of returns that may contain errors when claiming certain benefits (i.e. outcome (3) divided by outcome (2))?

To model these outcomes, we will use a regression of the following form:

(4) $$y_i = \zeta_0 + \zeta_1 ClientsSent6138_i + \varepsilon_i$$

Where y_i is the outcome for return preparer i and $ClientsSent6138_i$ adopts the value one if return preparer *i* had clients that were sent Letter 6138s, zero otherwise. In certain specifications, we may include a vector of return-preparer-level covariates.

Since outcomes (2) and (3) are both count variables, we will model these outcomes using ordinary least squares (assuming that the outcome is continuous), or as count outcomes using a Poisson or Negative Binomial model. This will depend on whether there is sufficient variation in the number of returns.

Since outcomes (3) and (4) are contingent on the return preparer remaining in the market, we will need to make adjustments for selective attrition. We will do this two ways. First, we will conduct a bounding exercise, imputing values for return preparers who drop out as either 0 or 1, and seeing the range of possible values for ζ_1. Second, we will predict attrition using return preparer covariates, and match treated return preparers who remained in the sample to untreated return preparers who have similar propensities to attrit. This is designed to re-balance the sample and allow for causal inference.

Link to an Analysis Code/Script:

N/A

Post-Commitment Adjustments