

IRS Preparer-Level Treatment Tests

Karen Masken, Return Preparer Office, Internal Revenue Service¹

Introduction

In 2010, the Internal Revenue Service (IRS) adopted regulations aimed at establishing standards for paid tax return preparers who prepare individual income tax returns (Form 1040). The objective was to improve voluntary compliance through increased oversight of the paid preparer industry with the goal of reducing errors on tax returns. The Return Preparer Office (RPO) was formed to meet this objective. RPO does not have enforcement authority, but rather focuses primarily on education, outreach, and partnering with the paid preparer community. Specifically, the three primary strategic goals of RPO are:

1. Register and promote a qualified tax professional community;
2. Improve the compliance and accuracy of returns prepared by tax professionals; and
3. Support a stakeholder-focused culture that encourages voluntary compliance and continuous improvement.

One effort to meet these strategic goals was a new requirement, effective January 1, 2011, that all paid preparers who prepare Form 1040 returns: register with the IRS; obtain a preparer tax identification number (PTIN); and enter the PTIN exclusively as the preparer identifying number on the returns they complete. Previously, PTINs had been optional.

Another effort involved a multi-year study that started in 2012 to determine the effect on tax return preparation accuracy of various treatments applied to paid tax return preparers. Because IRS has historically focused on taxpayer-level treatments, there is currently only a limited understanding of how preparer-level treatments affect change in preparer and client tax compliance. The goal of the multi-year study is to understand what treatments are effective on different segments of the noncompliant preparer population in an effort to determine the cost-effectiveness of the treatments. Details on the development and design of the study and results from the first year were presented in an earlier paper.² This paper builds upon the earlier version and includes additional information regarding the second year of the study.

Motivation

IRS resources have become increasingly scarce in the recent budget environment. The driver of the preparer-level treatment approach is that treatment of a single preparer is likely to improve the compliance of many tax returns, increasing the expected return on investment of treatment resources. This is similar to intervening at a wholesale level, rather than at the retail level.

As illustrated in Figure 1, preparer compliance is not a binomial variable, but rather a continuous spectrum. On one end of the spectrum are compliant, well-informed preparers, while on the other end of the spectrum are preparers who willfully perpetrate fraud. In the middle of the spectrum are those who might be unintentionally making errors due to lack of knowledge or those who are willfully noncompliant in their tax preparation but could potentially be moved toward voluntary compliance with a light touch. Moving preparers toward greater voluntary compliance is the most cost-effective action in that it protects revenue by having tax and credits reported correctly on the tax return—rather than IRS trying to recover revenue through examination and collection after the return has been filed, and in many cases, after a refund has been paid.

¹ The views expressed in this paper are those of the author and do not necessarily reflect the official position of the Internal Revenue Service.

² Masken, Karen C. "IRS Preparer-Level Treatment Tests: Results from the First Year of a Multi-Year Study," December 2014.

FIGURE 1. The Preparer Compliance Continuum

Traditionally, IRS has focused its enforcement resources toward the noncompliant and fraudulent end of the spectrum on what are predicted to be the largest, most substantial problems. For less substantial cases and for preparers in the middle of the spectrum, the tradition has been to rely on nontargeted services, primarily broad-based education efforts such as tax forums and webinars. Preparers subject to enforcement are generally subject to examination and in some cases, criminal investigation and prosecution. While these actions are necessary in some circumstances, they are very resource-intensive and costly. Audits of preparers typically involve auditing approximately 30 of their clients (for individual taxpayers, the current audit rate is less than two percent overall). If the audit of the preparer leads to litigation, then the Justice Department becomes involved and a very costly case has to be put together. Typically, the Justice Department handles less than fifty such cases each year. Since there are currently about 700,000 registered preparers, traditional enforcement efforts can reach only a small fraction of them.

Tax return preparers, like taxpayers, have certain rights. While there is a desire by some to take draconian actions against a preparer thought to be noncompliant (“Put them out of business!”, “Shame them publicly!”, etc...), many of these proposed ideas fail to recognize that, in addition to enforcement resource constraints, a preparer is entitled to many of the same rights afforded to individual taxpayers. By law, the IRS cannot share with the general public—or even with the preparer’s clients—the identity of a preparer under suspicion of filing noncompliant returns, as that would violate disclosure statutes and would not afford the preparer due process. The only time a preparer’s identity is disclosed is when a criminal or civil suit is brought against the preparer, at which point the information becomes a matter of public record. Unfortunately, since these are the cases that make headlines, it leaves many with the impression that ALL preparers are nefarious. While there are unequivocally bad actors in the community, RPO’s view is that these are the exception, not the rule, and many noncompliant preparers are in the middle of the compliance continuum.

While preparers toward the noncompliant/fraudulent end of the spectrum may require more expensive and intrusive treatments (e.g., audits/injunctions), finding effective, lower cost treatments for preparers in the middle of the spectrum could have a significant impact on revenue collected. Many preparers in the middle currently go untouched as traditional IRS examination approaches would not be cost-effective and the preparer may not partake in IRS educational services. Many of these preparers simply may not fully understand all the rules and may benefit from education as opposed to enforcement. Others might knowingly make errors believing they go unnoticed, but could be moved toward voluntary compliance by a touch lighter than traditional enforcement. If less expensive treatments are found to be effective, they would allow for a larger number of preparers to be treated and moved toward voluntary compliance by the same finite set of resources, thus protecting more revenue. Therefore, the focus of the treatments being tested is on preparers in the middle of the compliance spectrum. It should also be noted that a preparer is defined as an individual, not the firm or business where the preparer works (e.g., many well-known tax preparation businesses employ many individual preparers).

Issue Selection

The first step in the test development process was to identify which tax compliance issues to address. The following describes the rationale behind the issue choices.

The National Research Program (NRP) is an initiative at IRS that conducts audits on a random sample of individual taxpayers, thereby providing unbiased estimates of compliance for most line items on Form 1040 (individual income tax return). RPO analyzed the NRP results from Tax Years 2006 and 2007 and looked at the tax gap report based on the 2001 NRP study (the most current tax gap report available at the time) to inform the decision as to which issue to address with the treatments in the first year. Not only was the overall

magnitude of the compliance problem considered, but also to what extent returns by paid preparers contributed to noncompliance.

The issue chosen for the initial year of treatment tests was Schedule C net income. RPO found that approximately 75 percent of Schedule C returns in the NRP study were completed by paid preparers, and about 75 percent of those returns had errors. The overall contribution of misreported Schedule C income to the tax gap was approximately \$68 billion in Tax Year 2001, which was almost 30 percent of the individual income tax gap.

In the second year of the study, Schedule C net income continued to be addressed. In addition, RPO decided to address the Additional Child Tax Credit (ACTC), as well. The ACTC is a relatively new tax credit that was introduced in 1998. While RPO estimated from the NRP data that the misreported amount for ACTC was only about 2.5 percent of the Schedule C misreported amount, the ACTC is viewed as a potentially emerging compliance issue. Since approximately 65 percent of returns claiming the ACTC are prepared by paid preparers, it is an issue that may benefit significantly from an educational treatment at the preparer level.

Model Development and Selection Criteria

Schedule C

RPO inherited an established examination plan begun in 2010 that identified preparers for treatment and either subjected them to an Educational Visit or sent them a letter stating there were errors on their returns. The filters used to identify these preparers were developed by subject-matter experts based on their individual experience. In an audit³ by the Treasury Inspector General for Tax Administration (TIGTA), the filters were critiqued for not being data-driven and were identified as a management challenge. TIGTA recommended that a data-driven selection method be developed. RPO agreed with this recommendation, and error detection models using the National Research Program (NRP) data were developed for the Schedule C net income issue.

Because the NRP sampling rate overall is less than 0.01 percent, it is rare for more than one client per preparer to be selected for the random NRP audits. Therefore, the model was initially developed at the taxpayer level and then rolled up to the preparer-level selection criterion. The model was developed using data from the NRP Tax Years 2006 and 2007 studies and tested on the Tax Year 2008 study. In part, this was out of necessity as the Tax Year 2008 data were not yet available when the model was being developed. However, this approach does provide the benefit of assessing how robust the model is to choice of tax year. The model performed as expected on the test data.

After the model was developed, an outside expert performed an independent evaluation of the model and its development. The expert found the methodology was appropriate and that the model was effective. One technique employed to evaluate the model was a confusion matrix wherein the model was compared to a random draw. Based on the confusion matrix, the model was found to be more effective than a random draw by seven standard deviations.

The results from the first year of the study were not available when the second year was implemented. While no refinements to the model for the second year could be made based on new data, some refinements were made by employing the confusion matrix evaluation technique on subsets of the existing data.

To be eligible for the Schedule C test, preparers had to have an active PTIN, had to prepare at least 20 Schedule C returns with at least 15 percent of all their returns containing a Schedule C, and the majority (51 percent or more) of their Schedule C returns had to be flagged by the error-detection model. The requirement that they have at least 20 returns and 15 percent of their total returns with a Schedule C was simply to ensure that Schedule C was prevalent enough in the preparer's business to warrant treatment. Each of a preparer's Schedule C returns was scored using the respective model each year, and those considered high risk were flagged. The Schedule C returns were then aggregated by preparer and the percent of the preparer's Schedule C returns that were flagged was calculated. While it is virtually impossible to solve the endogeneity issue of

³ Implementation of the Return Preparer Visitation Project Was Successful, but Improvements Are Needed To Increase Its Effectiveness, June 29, 2012, TIGTA Audit # 2012-30-068.

whether it is the preparer or the client driving the noncompliance, RPO believed that if the majority of a preparer's Schedule C returns were flagged, then the preparer was likely to be at least partly responsible for the noncompliance. Therefore, if at least 51 percent of a preparer's Schedule C returns were flagged, then the preparer was placed in the treatment pool. There were preparers who met all the criteria except that they did not have an active PTIN. These preparers are both programmatically noncompliant (meaning they did not adhere to the basic requirement that they obtain a PTIN and enter it exclusively as the preparer identifying number on the returns they completed) and are at risk for tax noncompliance as well. RPO believed these preparers needed a different treatment to address both issues, so they were excluded from this set of tests.

ACTC

The initial ACTC analysis was based on information from the Return Transaction File for Calendar Year 2012. At the time the analysis was being performed, there was concern in Congress regarding the validity of children with Individual Tax Identification Numbers (ITINs) being claimed for the ACTC. The IRS issues ITINs to individuals who are required to have a taxpayer identification number but who do not have, and are not eligible to obtain, a Social Security Number. While children with ITINs can be claimed legitimately for the credit, they must meet specific criteria. One interesting phenomenon the data showed was that approximately 1 percent of return preparers accounted for 60 percent of all ACTC ITIN children. In light of this, two separate selection criteria were developed—one for preparers with a significant number of ACTC children with ITINs (dubbed 'ITIN Specialists') and another for preparers who had a significant number of ACTC returns, but not a significant number of ACTC children with ITINs (dubbed 'ACTC Generalists'). In contrast to the Schedule C model, the ACTC model is not an error detection model and does not attempt to make any inferences regarding the compliance of the preparers' returns.

The first set of eligibility criteria for ACTC mirrored those of Schedule C. The preparer had to have an active PTIN, prepare at least 20 ACTC returns, and at least 15 percent of the returns had to have ACTC claims.

To be eligible for the ITIN Specialist treatment, preparers had to have at least 20 returns that had children with ITINs being claimed for ACTC, and at least 15 percent of all their ACTC returns had to include children with ITINs being claimed for ACTC.

For preparers who did not meet the ITIN Specialist criteria, the percent of the preparer's returns that included ACTC was increased from 15 percent to at least 50 percent. This was simply to reduce the pool of eligible preparers and to focus on those with a greater prevalence of ACTC returns.

Treatments

In the initial year of the study, three different types of Schedule C treatments were tested and compared to one another. The first treatment was an "Educational Visit" to the preparers by revenue agents to discuss Schedule C issues found on returns that they had prepared. The second was a "Due Diligence Letter" reminding preparers of their due diligence requirements when preparing returns, and warning that they and their clients might be subject to audit (see Appendix A). The third treatment was a "Continuing Education Letter" with the same message regarding due diligence, but also recommending that, as part of the continuing education required at that time,⁴ the preparer take a minimum of 4 hours of continuing education regarding Schedule C (see Appendix B).

In the second year of the study, two of the three Schedule C treatments were re-employed and an additional two ACTC treatments were tested. For Schedule C, the Educational Visit and Due Diligence Letter tests were repeated. At the time, results from the first year were not in. However, since the first year of the tests, the IRS had lost the *Loving et al.* lawsuit and, as a result, could no longer require continuing education. While the Continuing Education Letter did not specifically state that a preparer *must* take continuing education, RPO

⁴ At the time of the first year of the test, preparers were required to take continuing education. It was later determined by the courts that IRS did not have the statutory authority to impose this requirement (*Loving et al. v. IRS*).

did not want to give the appearance that any type of education was required. Therefore, this treatment was not applied in the second year.

Two different letter treatments were tested for ACTC. In contrast to the Schedule C treatments, the ACTC treatments made no inference regarding the accuracy of the preparers' returns. Instead, they were purely educational and highlighted changes that had recently been made to ACTC filing requirements. Specifically, that what had previously been Schedule 8812 was now Form 8812, *Child Tax Credit*, and required to be filed with the return. One letter, the ITIN Specialist Letter, pointed out changes related to children with ITINs specifically (see Appendix C), while the ACTC Generalist Letter did not (see Appendix D).

Evaluation Method and Test Design

A Randomized Controlled Test (RCT) design was employed for each of the treatment tests, and the effectiveness of each treatment was evaluated using the Difference in Differences (DID) technique. Preparers in each of the treatment pools were randomly assigned to either the treatment or its respective control group. All treatments were applied prior to the start of the respective filing season (2013 the first year, and 2014 the subsequent year) and the determination of the effectiveness of each treatment was based on individual income tax returns filed in the respective years.

For the Schedule C treatments, the determination of effectiveness was based on the percent of clients who were flagged by the error detection model, described previously, in the filing season after the treatment was applied. During the design phase, it was decided that a treatment would be deemed successful for a particular preparer if it lowered the proportion of clients flagged for that preparer by 5 percentage points. While the definition of success will always be somewhat arbitrary, RPO believed it was important to define success during the design phase in order to avoid the perception that success was defined based on preliminary results (and the natural inclination to do so). The overall effectiveness for each treatment was determined by counting the number of successful preparers for each treatment and comparing it to the respective control group.

While preparers were randomly assigned to a treatment or control group, it is important to note that each treatment had a different set of constraints resulting in a different composition of preparers for each treatment and its respective control group. For the Educational Visits, the allocation of resources, namely Revenue Agents, had to be taken into consideration. Rather than drawing a simple random sample from the pool of preparers, a random sample proportionate to the resources available in each of seven IRS-defined geographic areas was drawn. For the Continuing Education Letter, preparers who held a credential (e.g., Certified Public Accountants and attorneys) were exempt from the IRS continuing education requirement in place at the time, as their own credential held them to a higher standard. They were, therefore, excluded from this treatment. The Due Diligence Letter had no constraints. Each of the treatments had its own control group with corresponding constraints; however, the control groups were not mutually exclusive of one another. Finally, each control group incorporated controls for possession of a credential and IRS area.

At the time the treatments were being implemented the first year, Hurricane Sandy struck land, and IRS generally suspends enforcement actions during a natural disaster. The visits were already underway, so preparers in the disaster area States (New Jersey, New York, Rhode Island, and Connecticut) were dropped from the test that year. Neither of the letters had been sent yet, so the test and control groups were redrawn excluding preparers in the disaster area. As a result, the size of the Educational Visit test in the first year is slightly smaller than the two letter tests.

Like the Schedule C tests, preparers in the ACTC tests were randomly assigned to either the test or control group for each treatment (note that, as described previously, the treatment groups are mutually exclusive). Since no error detection models were developed, the success of the tests was based simply on the number of ACTC claims the subsequent year, and for the ITIN Specialist, the number of children claimed for ACTC with ITINs.

Generally speaking, there are five statistical parameters that go into test-size calculations: precision, power, confidence level, the test statistic, and the critical value. For the Schedule C treatment tests specifically, the precision, or the probability of a false-positive (declaring success when it actually failed) was set at the standard

5 percent level. The power of the test, or the probability of a false-negative (declaring failure when it actually succeeded) was set at the standard level of 80 percent. The confidence level was set at 95 percent. The test statistic was the number of successes as defined above. The critical value was 5 percentage points, meaning that there had to be at least a 5 percentage point difference between the test and control to declare the difference statistically significant. The test size, given these parameters, was calculated to be approximately 1,250 preparers for each of the treatment and control groups.

Since the ACTC letters were educational in nature, there was an operational desire to send a larger number than a test would require. Therefore, the test size for each of these letters is larger than what was necessary statistically. There were 3,500 preparers in the ITIN Specialist Letter test and 5,000 in the ACTC Generalist Letter test.

At the initial design phase, an outside expert (different from the one employed to evaluate the model) provided consultation in the development of the test design to ensure the tests would produce the desired information.

Test Preparer Characteristics

Once the pools of preparers were identified, RPO worked with various other IRS offices to remove preparers who had been selected for enforcement efforts (such as criminal investigations). This undertaking is commonly referred to as 'deconfliction' within IRS. The final size of each test pool is presented in Table 1.

For the Schedule C tests, the final pool of preparers after deconfliction in year 1 was 9,600. This dropped in year 2 to 6,800. The main reason for the drop was that in addition to deconflicting with other IRS offices, preparers selected into either the test or control groups in the first year were excluded from the second year of the tests. Previously selected preparers account for 27 percentage points of the 39 percent dropped in year 2.

While the Schedule C tests lost only around 10 percent to deconfliction with other offices, the ACTC tests lost a much larger percentage. The primary reason for this was a simultaneous preparer-level treatment test by the IRS Earned Income Tax Credit (EITC) Program Office. Approximately 80 percent of all returns claiming ACTC also claim EITC, hence the large overlap. While RPO had not initially anticipated dropping so many preparers from the ACTC test pools, it turned out to be beneficial. While the ACTC preparers were not selected based on the risk of the returns they prepared, the preparers selected for EITC treatments were. As explained earlier in the paper, RPO's main focus is the preparers in the middle of the compliance spectrum who would not otherwise be touched. By removing preparers the EITC Program office believed to be high risk, the EITC Program Office helped RPO focus on preparers who are in the middle of the spectrum.

TABLE 1. Final Volume of Preparers Selected into Each Test Pool

Treatment Group	Number of Preparers Before and After Deconfliction		
	Before	After	% Dropped
Schedule C Year 1	10,600	9,600	9%
Schedule C Year 2	11,100	6,800	39%
ACTC ITIN Specialists	5,300	4,000	25%
ACTC Generalists	11,500	5,900	49%

On average, each preparer in the Schedule C pool in the first year prepared approximately 280 Form 1040 returns and 80 Schedule C returns (the medians were 180 and 50, respectively). For the second year, the average number of Form 1040 returns was 250 with an average of 70 Schedule C returns (with medians of 160 and 40, respectively). The average percent of their Schedule C returns flagged by the error detection model was 60 percent for both years.

In comparing the mean number of returns (280 in year 1 and 250 in year 2) and the median number of returns (180 in year 1 and 160 in year 2), it is clear that the volume of returns is highly skewed. The volume ranged from the required minimum of 20 returns to over 5,000 returns in each year. It is important to point

out that, in some instances, a preparer is under the supervision of another preparer who ultimately signs the return. Thus, while all the returns of the supervised preparer will bear that supervisor's PTIN, the supervisor is not necessarily the one who did the actual preparation. It is not possible to tell from the data whether a return was prepared by an unsupervised or supervised preparer.

For the preparers selected into the ACTC ITIN Specialist treatment pool, the average number of returns was 350, and 210 of those, on average, had ACTC claims. On average, 100 of their returns had at least one child with an ITIN being claimed for ACTC, with an average of 230 ACTC children with ITINs in total. The ACTC Generalist pool had an average of 100 returns and 60 ACTC claims. One note of interest is that the preparers who were excluded after deconfliction with the EITC Program Office tended to have a much higher volume of returns. The ITIN Specialists deconflicted had an average of 570 returns with 360 ACTC claims, and the ACTC Generalists had an average of 190 returns with 120 ACTC claims. Again, this illustrates that the final pool of preparers for the ACTC tests reflected preparers who would not normally be touched by IRS enforcement activities because they are too small to warrant enforcement treatment.

Results

This section begins with the general results of the Schedule C tests, followed by results of the ACTC tests. A more detailed analysis of the results for the first year of the Schedule C tests was presented in the earlier paper referred to previously, and is not repeated here.

Schedule C Tests

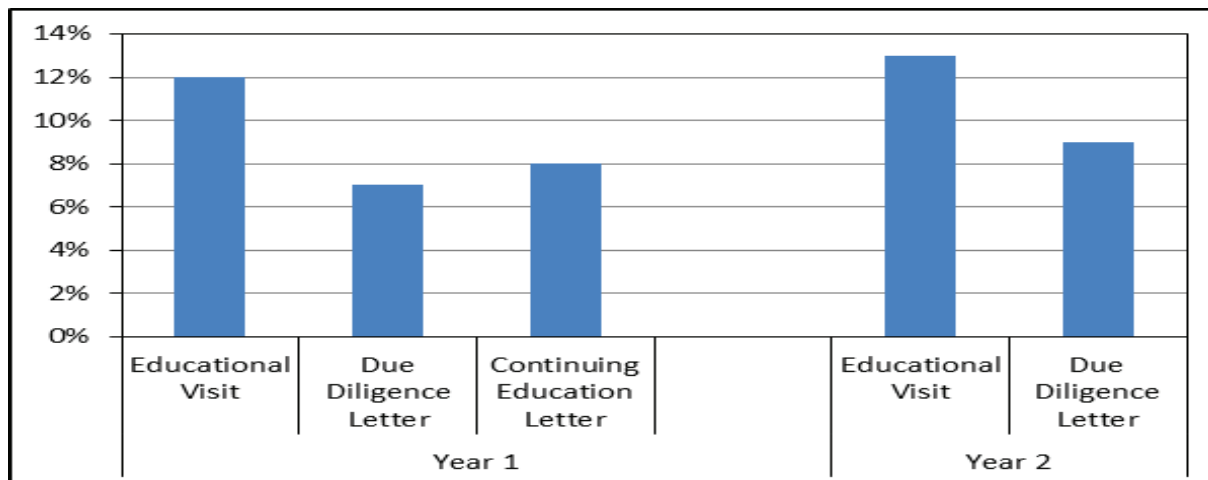
Attrition

One consequence of the treatments could be to effectively remove a preparer altogether from preparing returns. To the extent it put a noncompliant preparer out of business, this would be viewed as a positive result. On the other hand, it could also have an unintended negative consequence of moving the preparer to prepare but not sign returns, making them both programmatically and tax-reporting noncompliant. However, neither of these appears to have occurred to a significant degree.

For the Educational Visit treatment, the attrition rate for the test group in the first year was 3 percent, compared to 2 percent for the control group. However, this difference is not statistically significant. In the second year it was 2 percent for both the test and control groups. The preparers in the Due Diligence Letter test and control groups had a 2-percent attrition rate in both years. The Continuing Education Letter test had a 3-percent attrition rate for both the test and control groups. As previously discussed, the Continuing Education Letter test excluded CPAs and Attorneys, which may explain the slightly higher attrition rate for this group compared to the Due Diligence Letter treatment group since non-CPAs may be less invested in tax preparation.

Success Rate

As explained previously, the tests are evaluated by computing the difference in differences between the test and control groups. The differences, or success rates, for each of the Schedule C treatments are shown in Figure 2. As previously discussed, each of the treatment groups has a different composition and the test size for the Educational Visit treatment was smaller than the other two in the first year. Hence, the counts of success are not directly comparable to one another. While comparing the success rates is more accurate, it is still somewhat problematic in that each treatment had different constraints. However, if the treatments were to become operational, the constraints would remain the same. It is therefore beneficial to RPO to make the comparisons.

FIGURE 2. Success Rates of Schedule C Treatments

The difference in success rates between the test and control groups was highest for the Educational Visit treatment with a 12 percentage point difference between the test and control groups in the first year, and a 13 percentage point difference in the second year. In the first year, the Due Diligence Letter treatment had a success rate of 7 percentage points and the Continuing Education Letter treatment had a success rate of 8 percentage points. While these two are not statistically significantly different from one another, they are significantly lower than the Educational Visit success rate. As noted previously, in the second year the Continuing Education Letter treatment was dropped and the Due Diligence Letter was sent only to credentialed preparers. While the success rate for the Due Diligence Letter treatment in the first year was 7 percentage points overall, when broken down by credential, the noncredentialed preparers had a success rate of only 5 percentage points. Thus, the increase in year 2 to 9 percentage points is likely due to the fact that noncredentialed preparers were excluded from this treatment in the second year. Another explanation for the slightly higher rate for both the Educational Visit and the Due Diligence Letter treatments is that the underlying selection model was improved.

While the Educational Visit treatment was significantly more effective, it is not as cost-effective as either of the letter treatments. It is difficult to get a firm comprehensive cost estimate of a visit. However, the order of magnitude is really all that is necessary. The cost of an Educational Visit is over \$100, while the cost of a letter is under \$1. Thus, the effect of the Educational Visit would need to be much higher than it is in order for it to be cost-effective.

For preparers where the Continuing Education Letter treatment was considered successful, RPO was interested in what percentage took continuing education as a result of the letter. Currently, data available to RPO do not show the topic or the actual date of the continuing education, but there was a 5-percentage point difference in the percent of preparers who took continuing education between the test group (31 percent) and control group (26 percent). Hence, it appears the letter did effectively encourage preparers to take at least some form of continuing education.

Overall, the successful preparers in the first year prepared more than 375,000 individual tax returns with around 100,000 returns including a Schedule C. While a success was defined as a 5-percentage point drop in the number of Schedules C flagged by the error-detection model, the actual average drop across the treatments was 18 percentage points. The estimated number of taxpayers who moved towards voluntary compliance as a result of the three treatments in year 1 is approximately 18,300 taxpayers.

For the second year, successful preparers prepared more than 315,000 individual returns with over 80,000 Schedules C. As in the first year, the average percentage of Schedule C returns flagged by the error detection model dropped by 18 percentage points. The estimated number of taxpayers who moved towards voluntary compliance in the second year was approximately 14,400.

Client Migration

While preparers were the primary focus of the treatments, RPO was also interested in what happened with their clients. As mentioned earlier, there is an issue of endogeneity when trying to determine if the taxpayer or the preparer is driving the noncompliance. To the extent it is the taxpayer driving the noncompliance, then for successfully treated preparers, one might expect these clients to discontinue using that preparer and either go to a new preparer or prepare the return themselves. Clients of successfully treated preparers who were flagged by the error-detection model in both years actually had the same or lower migration rates in the test group than the control group for all three treatments. For the Educational Visit, the test group rate was 22 percent compared to the control group rate of 26 percent. For the Due Diligence Letter, the rate was 27 percent for both the test and control groups. Finally, the rates for the Continuing Education Letter were 30 percent and 33 percent for the test and control groups, respectively. Thus, it appears that the model was effective in identifying preparers who were driving the noncompliance.

Recidivism

While all three treatments were found to be effective for the filing season immediately following the treatment, RPO wanted to know if the effect was persistent in the following year, as well. The recidivism issue was addressed by comparing the success rate 2 years later for preparers who were successful in year 1. In comparing the success rate in year 2 between the test and control groups, there was no statistically significant difference between the test and control for any of the treatments, and therefore recidivism does not appear to be an issue. For the Educational Visit treatment, the success rate for the test was 76 percent, compared to 70 percent for the control. The success rates for the Due Diligence Letter and the Continuing Education Letter test groups were 68 percent and 71 percent, respectively, compared to 70 percent for both control groups.

ACTC Tests

Attrition

Like the Schedule C treatments tests, the ACTC Treatment tests had no effect on the attrition rate for preparers. The attrition rate for the ITIN Specialists was 6 percent for both the test and control groups. The ACTC Generalist preparers were much more volatile with rates of 25 percent for the control group and 26 percent for the test group, but these rates were not statistically significantly different from one another.

Success Rates

As stated previously, both of the ACTC treatments were purely educational in nature and no inference regarding the preparers' compliance was made during the selection process. However, if the treatments effectively move preparers towards voluntary compliance, this can be inferred by simply comparing the change in the number of claims between the test and control groups. Additionally, for the ITIN Specialist Letter treatment a comparison between the test and control group of the number of children with ITINs being claimed provides an indication of the effectiveness of the educational letter.

The average number of ACTC claims for preparers in the ACTC Generalist Letter test group was 4-percentage points lower than the control group. The average total amount of ACTC claimed was also 4-percentage points lower than the control group. The average number of ACTC claims for preparers in the ITIN Specialist Letter test group was also 4-percentage points lower than the respective control group, and the average amount of the ACTC claims was 5-percentage points lower. In addition, the average total number of children with ITINs claimed was 10-percentage points lower than the control group. All of these differences are statistically significant; therefore, all treatments were effective in reducing the number of ACTC claims and moving the preparers towards voluntary compliance.

Preparers in the ACTC Generalist Letter treatment test group prepared more than 480,000 returns in the selection year with more than 285,000 ACTC claims. As a result of the treatment there were an estimated 2,500 fewer ACTC claims. For the ITIN Specialist Letter treatment, preparers had approximately 745,000 ACTC claims with about 800,000 children with ITINs on 1.2 million individual returns. The treatment resulted in an

estimated reduction of 1,800 ACTC claims overall, but a reduction of about 18,700 children with ITINs being claimed on ACTC returns.

Client migration

As with the Schedule C tests, there was some concern that if a preparer told a client that the client was not eligible for ACTC, that client might simply move to a different preparer who would make the claim. In following the clients however, this does not appear to be an issue. For the ACTC Generalist Letter treatment, 48 percent of the clients claiming ACTC the first year stayed with the same preparer, compared to 47 percent of clients for the control group. For the ITIN Specialist Letter treatment, 34 percent of clients claiming ACTC in the first year stayed with the same preparer, compared to 31 percent of the control group.

An additional concern for the ITIN Specialist group was that the child with an ITIN would simply be claimed by someone else on a different tax return. However, the attrition rate for the children with ITINs being claimed by anyone was higher for the test group than the control group (31 percent and 28 percent, respectively), indicating that child migration was not a significant issue.

Summary

Employing a statistically valid Randomized Controlled Test and the Difference in Differences evaluation technique allowed RPO to learn the effectiveness of each of the treatments and make appropriate revisions. The results from the first year of the Schedule C tests indicate all three treatments were effective. The second year indicated that the results were consistent from year to year, and that results from the first year were persistent in the following year. However, while the Educational Visit treatment was more effective than either letter, it is not as cost-effective. As a result of the test, and due to budget constraints within the IRS, the Educational Visit treatment was dropped from the treatment pool in subsequent years. The ACTC tests demonstrated that treatments that are purely educational in nature can also be effective. These educational letters are more cost-effective than the targeted compliance letters since no underlying model needs to be developed to select preparers for treatment. As a result, RPO is currently testing an educational Schedule C treatment in lieu of the target compliance letters, as well as continuing the ACTC letters.

In summary, the initial tests were designed to target specific preparers for compliance treatments. However, the results from the first 2 years of the tests have moved RPO away from targeted treatments and towards more general and less costly educational efforts.

References

- Brown, Robert E., and Mark J. Mazur. (2013) "IRS's Comprehensive Approach to Compliance Measurement," Presentation to the National Tax Association Spring Symposium, Washington, DC.
- Hastie, Trevor; Robert Tibshirani; and Jerome Friedman. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science+Business Media, LLC, New York, NY.
- Hicks, Charles R. (1993) *Fundamental Concepts in the Design of Experiments*. Oxford University Press, New York, NY.
- Internal Revenue Service Data Book*, Fiscal Year 2012. (2013) Publication 55B, Washington, DC.
- Masken, Karen C. (2014) "IRS Preparer-Level Treatment Tests: Results from the First Year of a Multi-Year Study." http://www.irs.gov/pub/irs-utl/irs_preparer_level_treatment_tests.pdf.
- Mazur, Mark J., and Alan H. Plumley. (2007) "Understanding the Tax Gap," Presentation to the National Tax Association Spring Symposium, Washington, DC.
- Treasury Inspector General for Tax Administration. (2012) "Implementation of the Return Preparer Visitation Project Was Successful, but Improvements Are Needed To Increase Its Effectiveness." Report 2012-30-068, June 29, 2012.

Appendix A

Due Diligence Letter



Department of the Treasury
Internal Revenue Service
Return Preparer Office
1122 Town and Country Commons
Chesterfield, MO 63017

Letter	XXXX
Date	Date
To contact us	Phone 1-636-255-1208 8 a.m.- 5 p.m. CT

Tax Return Preparer
Address
City, ST zip

Subject: Reminders about Schedule C preparation

Dear Tax Return Preparer,

A review of tax returns you have prepared in the past year shows that many have a high percentage of traits we believe typically indicate errors in preparing Form 1040, Schedule C, *Profit or Loss from Business (Sole Proprietorship)*. This letter is to remind you of your responsibilities in this area, provide educational assistance, and request that you pay special attention to it next filing season.

Due diligence responsibilities

A paid tax return preparer is expected to take multiple steps to prepare accurate tax returns on behalf of clients. These include reviewing the applicable tax law, and establishing the relevancy and reasonableness of income, credits, expenses, and deductions to be reported on the return. In general, you may rely in good faith without verification upon information furnished by the client. However, you may not ignore the implications of information furnished to, or actually known by you. You must make reasonable inquiries if the information appears to be incorrect, inconsistent with an important fact or another factual assumption, or incomplete.

Schedule C reminders

To prepare accurate Schedules C, you should ask your clients sufficient questions to determine that the expenses claimed are correct and allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they are entitled to claim deductions for nonqualifying expenditures. You should also ask your clients if they have documentation to support the expenses in case receipts are requested by the IRS.

Helpful resources

Based on our analysis of the Schedules C you prepared, we encourage you to review the Schedule C instructions and other IRS publications available at www.irs.gov, keyword: Recommended Reading for Small Businesses. We also encourage you to review Circular 230, Regulations Governing Practice Before the Internal Revenue Service, sections 10.22 and 10.34, titled "Diligence as to accuracy" and "Standards with respect to tax returns and documents, affidavits, and other papers," respectively.

Potential consequences

In the future, both you and your clients may be adversely affected by incorrect returns. Consequences may include any or all of the following:

- If your clients' returns are examined and found to be incorrect, your clients may be liable for additional tax, interest, and penalties.

- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to an unreasonable position can be assessed a penalty of at least \$1,000 per return (IRC section 6694(a)).
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to reckless or intentional disregard of rules or regulations by the tax preparer, can be assessed a penalty of at least \$5,000 per return (IRC section 6694(b)).

We hope this letter has heightened your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Schedules C for your clients.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

Appendix B

Continuing Education Letter



Department of the Treasury
Internal Revenue Service
Return Preparer Office
1122 Town and Country Commons
Chesterfield, MO 63017

Letter	XXXX
Date	Date
To contact us	Phone 1-636-255-1208 8 a.m.- 5 p.m. CT

Tax Return Preparer
Address
City, ST zip

Subject: Recommendation to take CE programs about Schedule C

Dear Tax Return Preparer,

A review of tax returns you have prepared in the past year shows that many have a high percentage of traits we believe typically indicate errors in preparing Form 1040, Schedule C, *Profit or Loss from Business (Sole Proprietorship)*. Therefore you may benefit from continuing education (CE) programs on this topic.

The purpose of this letter is to recommend that as part of your 2013 CE requirement, you take a **minimum of 4 hours** of programs related to business income and/or expenses. Information about CE requirements and a list of IRS approved CE providers is available at www.irs.gov/taxpros/ce.

In addition, as you prepare returns for the next filing season, please pay special attention to your work on Schedule C returns to ensure they are prepared accurately. In particular, make sure you have familiarized yourself with the following areas:

Due diligence responsibilities

A paid tax return preparer is expected to take multiple steps to prepare accurate tax returns on behalf of clients. These include reviewing the applicable tax law, and establishing the relevancy and reasonableness of income, credits, expenses, and deductions to be reported on the return. In general, you may rely in good faith without verification upon information furnished by the client. However, you may not ignore the implications of information furnished to, or actually known by you. You must make reasonable inquiries if the information appears to be incorrect, inconsistent with an important fact or another factual assumption, or incomplete.

Schedule C reminders

To prepare accurate Schedules C, you should ask your clients sufficient questions to determine that the expenses claimed are correct and allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they are entitled to claim deductions for nonqualifying expenditures. You should also ask your clients if they have documentation to support the expenses in case receipts are requested by the IRS.

Helpful resources

Based on our analysis of the Schedules C you prepared, in addition to considering CE programs, we encourage you to review the Schedule C instructions and other IRS publications available at www.irs.gov, keyword: Recommended Reading for Small Businesses. We also encourage you to review Circular 230, Regulations Governing Practice Before the Internal Revenue Service, sections 10.22 and 10.34, titled "Diligence as to accuracy" and "Standards with respect to tax returns and documents, affidavits, and other papers," respectively.

Potential consequences

In the future, both you and your clients may be adversely affected by incorrect returns. We will check whether you complete the continuing education on business income and/or expenses as recommended and we will be looking for improvements in returns you prepare. Incorrect returns may cause any of the following consequences:

- If your clients' returns are examined and found to be incorrect, your clients may be liable for additional tax, interest, additions to tax, and penalties.
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to an unreasonable position can be assessed a penalty of at least \$1,000 per return (IRC section 6694(a)).
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to reckless or intentional disregard of rules or regulations by the tax preparer, can be assessed a penalty of at least \$5,000 per return (IRC section 6694(b)).

We hope this letter and your continuing education focus on Schedule C preparation will heighten your awareness of your responsibilities and help ensure you prepare accurate Schedules C for your clients.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

Appendix C

ACTC ITIN Specialist Letter

Tax Return Preparer
Address
City, ST zip

Dear Tax Return Preparer,

Our information indicates that you prepared income tax returns for Tax Year 2012 claiming the Child Tax Credit and Additional Child Tax Credit. The procedures for claiming these credits were modified for 2012 with the creation of Form 1040, Schedule 8812, *Child Tax Credit*. As a paid preparer, you are responsible for ensuring your clients' returns are accurate.

The purpose of this letter is to provide educational assistance and promote awareness of the new Schedule 8812 requirements. Schedule 8812 has four parts and your clients' circumstances will determine which parts must be completed. We ask that you pay special attention to the requirements for each part, including the requirements for claiming the Child Tax Credit and Additional Child Tax Credit for children with an Individual Taxpayer Identification Number (ITIN) next filing season.

ITIN and Child Tax Credit reminders

To prepare an accurate Schedule 8812, you must ask your clients relevant and probing questions to help you determine if the credit is allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they can claim the credit for ineligible dependents. In general, to be a qualifying child for purposes of the child tax credit, the child must be a citizen, national, or resident of the United States. You should ask your clients questions to identify whether a child with an ITIN meets the substantial presence test for establishing residency, if the child is not a U.S. citizen or U.S. national.

Due diligence responsibilities

A paid tax return preparer must take multiple steps to prepare accurate tax returns on behalf of clients. These steps are a preparer's due diligence and include reviewing the applicable tax law to establish the relevance and reasonableness of income, credits, expenses, and deductions on a return. In general, you can rely in good faith without verification on information your client provides. However, you can't ignore the implication of the information you have. You must make reasonable inquiries if the information appears to be incorrect, inconsistent, or incomplete.

Helpful resources

Specific information about the Schedule 8812 is available on our website at www.irs.gov, keyword: Child Tax Credit. In addition, we recommend you review:

- Schedule 8812 instructions
- Circular 230, Section 10.22, Diligence as to accuracy
- Circular 230, Section 10.34, Standards with respect to tax returns and documents, affidavits, and other papers

I hope this letter has increased your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Child Tax Credit claims for your clients.

Thank you for your attention to this matter.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office

Appendix D

ACTC Generalist Letter

Tax Return Preparer
Address
City, ST zip

Dear Tax Return Preparer,

Our information indicates that you prepared income tax returns for Tax Year 2012 claiming the Child Tax Credit and Additional Child Tax Credit. The procedures for claiming these credits were modified for 2012 with the creation of Form 1040, Schedule 8812, *Child Tax Credit*. As a paid preparer, you are responsible for ensuring your clients' returns are accurate.

The purpose of this letter is to provide educational assistance and promote awareness of the new Schedule 8812 requirements. Schedule 8812 has four parts and your clients' circumstances will determine which parts must be completed. We ask that you pay special attention to the requirements for each part next filing season.

Child Tax Credit reminders

To prepare an accurate Schedule 8812, you must ask your clients relevant and probing questions to help you determine if the credit is allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they can claim the credit for ineligible dependents. In general, to be a qualifying child for purposes of the child tax credit, the child must be a citizen, national, or resident of the United States. You should ask your clients questions to identify whether a child meets the substantial presence test for establishing residency, if the child is not a U.S. citizen or U.S. national.

Due diligence responsibilities

A paid tax return preparer must take multiple steps to prepare accurate tax returns on behalf of clients. These steps are a preparer's due diligence and include reviewing the applicable tax law to establish the relevance and reasonableness of income, credits, expenses, and deductions on a return. In general, you can rely in good faith without verification on information your client provides. However, you can't ignore the implication of the information you have. You must make reasonable inquiries if the information appears to be incorrect, inconsistent, or incomplete.

Helpful resources

Specific information about the Schedule 8812 is available on our website at www.irs.gov, keyword: Child Tax Credit. In addition, we recommend you review:

- Schedule 8812 instructions
- Circular 230, Section 10.22, *Diligence as to accuracy*
- Circular 230, Section 10.34, *Standards with respect to tax returns and documents, affidavits, and other papers*

I hope this letter has increased your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Child Tax Credit claims for your clients.

Thank you for your attention to this matter.

Sincerely,

Carol A. Campbell
Director, Return Preparer Office